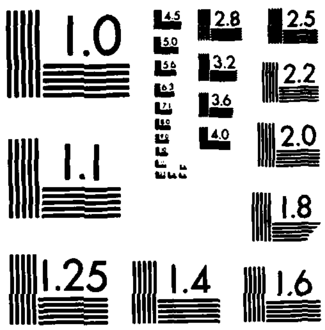
AD-A155 100 AN EVALUATION OF DISCRETIZED CONDITIONAL PROBABILITY AND LINEAR REGRESSIO..(U) NAVAL POSTGRADUATE SCHOOL MONTEREY CA M H WOOSTER DEC 84 1/2

UNCLASSIFIED F/G 4/2 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

AN EVALUATION OF DISCRETIZED CONDITIONAL PROBABILITY AND LINEAR REGRESSION THRESHOLD TECHNIQUES IN MODEL OUTPUT STATISTICS FORECASTING OF CLOUD AMOUNT AND CEILING OVER THE NORTH ATLANTIC OCEAN

by

Michael H. Wooster

December 1984

Thesis Advisor: Robert J. Renard

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO. AD-A155100

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
An Evaluation of Discretized Conditional Probability and Linear Regression Threshold Techniques in Model Output Statistics Forecasting of Cloud Amount and Ceiling Over the North Atlantic Ocean

5. TYPE OF REPORT & PERIOD COVERED
Master's Thesis
December 1984

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
Michael H. Wooster

8. CONTRACT OR GRANT NUMBER(s)

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Naval Postgraduate School
Monterey, California 93943

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Naval Postgraduate School
Monterey, California 93943

12. REPORT DATE
December 1984

13. NUMBER OF PAGES
187

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Model Output Statistics, Cloud Amount, Ceiling, Discretization, Conditional Probabilities, Categorical Forecasting, Ocean Areas, Thresholds, Linear Regression, Natural Regression, Maximum Probability, Measures of Separability, Cluster Analysis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report describes the application and evaluation of several statistical models in the forecasting of cloud amount and ceiling over selected physically homogeneous areas of the North Atlantic Ocean. The focus of this study is to evaluate the applicability of previous Naval Postgraduate School model output statistics research in the area of horizontal marine visibility to the forecasting of cloud amount and ceiling over

#20 - ABSTRACT - (CONTINUED)

ocean areas. The models, including minimum probable error linear regression threshold techniques, maximum conditional probability and natural regression utilize observed visibility data and model output parameters from the Navy Operational Global Atmospheric Prediction System (NOGAPS). Results show statistically similar results for the linear regression and maximum conditional probability models. Also included is the result of additional experimentation on the application of several measures of separability and cluster analysis to predictor selection.

Approved for public release; distribution unlimited.

An Evaluation of Discretized Conditional Probability and Linear Regression Threshold Techniques in Model Output Statistics Forecasting of Cloud Amount and Ceiling Over the North Atlantic Ocean

by

Michael H. Wooster
Lieutenant Commander, United States Navy
B.S., United States Naval Academy, 1975

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN METEOROLOGY AND OCEANOGRAPHY

from the

NAVAL POSTGRADUATE SCHOOL
December 1984

Author: Michael H. Wooster
Michael H. Wooster

Approved by: Robert J. Renard
Robert J. Renard, Thesis Advisor

Paul R. Lowe
Paul R. Lowe, Second Reader

Robert J. Renard
Robert J. Renard, Chairman,
Department of Meteorology

J.N. Dyer
John Dyer,
Dean of Science and Engineering

ABSTRACT

This ~~report~~ describes the application and evaluation of several statistical models in the forecasting of cloud amount and ceiling over ~~selected physically homogeneous~~ areas of the North Atlantic Ocean. The focus of this study is to evaluate the applicability of previous Naval Postgraduate School model output statistics research in the area of horizontal marine visibility to the forecasting of cloud amount and ceiling over ocean areas. The models, including minimum probable error linear regression threshold techniques, maximum conditional probability and natural regression, utilize observed visibility data and model output parameters from the Navy Operational Global Atmospheric Prediction System (NOGAPS). Results show statistically similar results for the linear regression and maximum conditional probability models. Also included is the result of additional experimentation on the application of several measures of separability and cluster analysis to predictor selection.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

IV. PROCEDURES

A. TERMS AND SYMBOLS

The terms and statistical symbols defined below will be used throughout the remainder of this report. The formal mathematical definitions are described in Karl (1984).

1. Maximum probability strategy--choosing the forecast weather element (e.g., cloud amount or ceiling) category based upon the highest probability of the weather element within a predictor interval, hence conditional probability.

    a. MAXPROB I--designation of the maximum probability strategy in which ties of the highest conditional probabilities in a predictor interval are resolved by the generation of a random number.

    b. MAXPROB II--designation of the maximum probability strategy in which ties of the highest conditional probabilities in a predictor interval are resolved by assigning the lowest element category, of those tied, as the forecast category.

2. Natural regression strategy--choosing weather categories based upon the statistical average of the conditional probabilities of the weather element within a predictor interval.

3. A0--the probability of a zero-class weather element category forecast error (e.g., if cloud amount category I is forecast and observed). This is more generally known as total percentage correct.

4. A1--the probability of a one-class weather element category forecast error (e.g., if ceiling category I is forecast and category II is observed).

5. A2--the probability of a two-class weather element category forecast error (e.g., if ceiling category I is forecast and category III is observed).

6. CE--class error parameter defined as A1+2A2, used as the primary aid in identifying the first predictor for the Preisendorfer strategies.

D. TRAINING/TESTING DATA SETS

One-third of the observations were withheld from the developmental model to use as an independent data set (the testing set). This was accomplished by the use of a counter and transfer statement in the computer programs which prevented every third observation from entering the developmental computations. Although the approach has the advantage of simplicity, there could be some sort of ordering in the data base, hence the split runs a chance of being non-random. To ensure that the dependent (the training data set) and independent (the testing data set) data were representative of the same population, a 95% confidence interval for proportions (Miller and Freund, 1977) was established from the entire data set, for each of the weather element categories; the training and testing data sets were constrained to have frequencies of occurrence within these established confidence intervals.

co-located with the National Climatic Data Center (NCDC). The observations which were obviously erroneous, as determined from the data quality indicators provided with the data, were deleted from the working data sets.

4. Predictor Parameters

Fifty TAU-00, fifty-four TAU-24 and fifty-four TAU-48 model output predictors (MOP's) were provided by the Fleet Numerical Oceanography Center (FNOC), Monterey, California. These parameters are generated by their current operational atmospheric prediction model, the Navy Operational Global Atmospheric Prediction System (NOGAPS). All MOP's were interpolated from model grid coordinates to synoptic ship report position using a linear interpolation scheme. In addition to the initial group of model output parameters, ten derived parameters representing calculated quantities, such as parameter gradients, products and advections, were included as potential predictors. A listing of all available TAU-00, TAU-24 and TAU-48 MOP's are included in Appendix D.

For each homogeneous area and model forecast projection, a set of two linear regression equations, in addition to the aforementioned MOP's, were included as potential MOP's for a separate evaluation of the Preisendorfer methodology (the PR+BMD model). These two predictor equations were obtained from a standardized linear regression software package, P9R, an all possible subsets regression, as addressed in the BMDP Statistical Software (University of California, 1983).

2. Time Period

Data from mid-May 1983 to mid-July 1983 were combined to form a more extensive data set, hereafter referred to as FATJUNE 1983. FATJUNE 1983 was selected as the initial data set for the visibility studies due to its high frequency of occurrence of poor visibility observations, and it was chosen for this study to maintain continuity on the overall MOS project. 1200 GMT synoptic ship report data were used exclusively in this study since 1200 GMT corresponds to general daylight conditions over the North Atlantic Ocean during FATJUNE. For the purpose of this study, TAU-00 model output parameters (MOP) generally represent six-hour model forecasts valid at 1200 GMT. However, three specific fields, namely temperature, geopotential height and wind, are model initialization fields at 1200 GMT. TAU-24 and TAU-48 MOP's are 24-hour and 48-hour model forecasts, respectively, valid at 1200 GMT. TAU-00, TAU-24 and TAU-48 MOP's (predictors) are employed in the 00-, 24- and 48-h forecast schemes, respectively. Summaries of the cloud amount and ceiling frequencies for each category type, as a function of homogeneous area and prediction time for FATJUNE 1983, are contained in Tables I through IV, respectively.

3. Synoptic Weather Reports

All synoptic weather observations (predictand data) for this study were provided by the Naval Oceanography Command Detachment (NOCD), Asheville, North Carolina which is

Ceiling Category	Code	Definition
I	0-3	< 1000 feet
II	4-5	1000-3500 feet
III	6-9	> 3500 feet

The above scheme is based on U.S. Navy operational criteria:

1. Ceiling less than 1000 feet--U.S. Navy aircraft carrier at-sea flight recovery operations require controlled (IFR) approach guidelines (Department of the Navy, 1979).
2. Ceiling 1000-3000 feet--flight recovery operations require modified IFR approach guidelines.
3. Ceiling greater than 3000 feet--at-sea recovery operations change to visual (VFR) approach guidelines.

C. NORTH ATLANTIC OCEAN DATA

1. Area

The North Atlantic Ocean, from 0° to 80°N latitude, was divided into homogeneous oceanic areas following Lowe (1984b), using a statistical cluster analysis technique. The specific homogeneous areas evaluated in this study are identified as areas 2 and 4 on Fig. 2. These areas were selected because they contain the largest data samples and represent two different relative frequencies of cloud cover. Area 4 represents an area where the three categories are of near equal population while area 2 represents an area where the number of category I (clear and scattered) observations is about one-half of those in categories two and three (broken and overcast).

as thin or partial. However, the synoptic Surface Marine Observations data set does not give a direct observation of ceiling height, thus making it necessary for the purpose of this study to synthesize ceiling height from the data that are given. The data set gives the following reported fields:

CLAMT: Cloud amount or total sky cover

LOAMT: Total sky cover by low clouds (middle clouds if no low clouds are present)

CLHT: Height of the lowest clouds irrespective of amount.

Cloud height is reported with a synoptic code from 0 to 9, where 0 refers to heights from 0 to 50 feet and 9 refers to heights greater than 6500 feet or cases where no clouds are present. This 6500 foot height, corresponds roughly to the upper boundary of the clouds reported in the low cloud amount field, LOAMT, making possible the following definition of ceiling for calculational purposes in this study:

If the reported LOAMT $<$ 5/8 then the ceiling is unlimited.

If the reported LOAMT $\geq$ 5/8 then ceiling is taken as the reported cloud height.

The ceiling observations are likewise treated as categorized predictands and are divided into the following categories for prediction purposes:

III. DATA

A. CLOUD AMOUNT OBSERVATIONS AND SYNOPTIC CODES

Cloud amount is defined as the fraction of the celestial dome covered by all clouds. The observations taken from seagoing platforms are reported as values of zero to eight oktas (eighths) such that 0 means no clouds, 1 means 1/8th cloud cover, etc. In addition, 9 is used to report an obscured sky (e.g., smoke, fog), for which a defined cloud cover is not observable. The observations were treated as categorized predictands and were divided into categories conforming to the standard definitions of opaque sky cover for clear, scattered, broken and overcast, as used for aviation observations.

Cloud Amount Category	Eighths	Definition
I	0-4	clear/scattered
II	5-7	broken
III	8	overcast

The obscured observation was not used in the MOS development reported on here.

B. CEILING OBSERVATIONS AND SYNOPTIC CODES

The definition of ceiling is the height ascribed to the lowest layer of clouds or obscuring phenomena when it is reported as broken, overcast, or obscured and not classified

II. OBJECTIVES AND APPROACH

The objective of this study is to extend the previous NPS research for predicting horizontal marine visibility using model output statistics (MOS) (Karl, 1984; Diunizio, 1984) to the prediction of cloud ceiling and cloud cover over coastal and open ocean areas of the North Atlantic Ocean. The approach to the problem is as follows:

A. Define categorical groupings of cloud amount and ceiling height which relate to operational use at sea.

B. Determine if one element is important to the prediction of the other and, therefore, should be investigated first.

C. Apply the previously investigated methods for forecasting visibility to cloud amount and ceiling, and evaluate their performance. These methods include Preisendorfer (1983 a,b,c) maximum probability and natural regression strategies, and linear regression threshold models as proposed by Lowe (1984a).

D. Compare and contrast the results of the two methodologies in C, above, and conduct some experimentation to improve their applicability to cloud cover and ceiling prediction.

E. Investigate alternative predictor selection schemes to improve the ability of the MOS models to distinguish between the predictand categories.

F. Make recommendations on the usefulness of the schemes investigated and potential avenues for future work in this area.

range of atmospheric visibility. Secondly, the number of observed weather reports are not only limited in number, but also come from moving platforms so that there is a lack of weather trend information for a single (fixed) station as is the case in land observations. Statistical methodologies tested by Karl (1984) and Diunizio (1984) to overcome the at-sea MOS problems, include a conditional probability approach proposed by Preisendorfer (1983 a,b,c) and various innovative threshold techniques, as applied to the linear regression model, developed by Lowe (1984a).

This study represents a continuation of the North Atlantic Ocean MOS studies on visibility, by Karl (1984) and Diunizio (1984). However, in this case, the statistical methods tested by the earlier visibility studies are applied to cloud amount and ceiling. The methods used here have been designed to be consistent with those of the previous studies in order to allow for comparison of results, as appropriate.

Prediction Research Facility (NEPERF) in Monterey, California sponsored a limited amount of research into naval applications of MOS, with most of the effort going toward marine visibility and fog. The results of these Navy studies and the encouraging performances of the NWS and AWS MOS programs prompted the Navy, in the spring of 1983, to begin development of a MOS program under the guidance of NEPERF, to forecast operational air/ocean parameters over the oceans of the world. The proposed milestones of this ten year project are summarized in Fig. 1. The first operational weather parameter investigated in the program is horizontal visibility over the North Atlantic Ocean using MOP's from the Navy Operational Global Atmospheric Prediction System (NOGAPS), a dynamical primitive equation (PE) model run operationally at the Fleet Numerical Oceanography Center (FNOC) (Karl, 1984; Diunizio, 1984).

Previous experimental work by the Navy to forecast open-ocean fog and visibility using linear regression equations (Aldinger, 1979; Yavorsky, 1980; Selsor, 1980; Koziara et al. 1983; Renard and Thompson, 1984) shows skill of marginal operational usefulness but exceeding that of persistence and/or climatology. Two factors limit the potential for MOS forecasts of visiblity and fog at sea. First, there is the lack of 'calibrated' fog and visiblity observations in that shipboard weather observers lack sufficient reference points to be able to accurately estimate the

assembled for a land station or region for a period of several years, stratified by season or month.

The National Weather Service (NWS) has included MOS as an integral part of their weather forecasting operations since the mid 1970's and currently forecasts for approximately 15 weather elements at forecasting times of 6 to 48 hours. These MOS forecast equations, developed by the National Oceanic and Atmospheric Administration's (NOAA) Techniques Development Laboratory (TDL), are based on model output parameters (MOP's) from the U.S. regional model, LFM-II. In December of 1980, the Air Force Air Weather Service (AWS) also implemented and operated a MOS forecasting scheme at the Air Force Global Weather Center (AFGWC), Offut AFB, Nebraska (Best and Pryor, 1983) for approximately 18 months. The program was terminated with the decision to replace their hemispheric primitive equation model with a spectral global dynamic model (Klein, 1981). The linear regression techniques used by both the Air Force and NWS has demonstrated operationally useful skill in forecasting weather elements at locations over land throughout the world (Best and Pryor, 1983). In this technique, called Regression Estimated Event Probability (REEP), predictor variables are discretized into sets of dummy variables prior to regression.

The Navy's unique responsibilities of marine forecasting provides a motive for it to have its own MOS system. In the late 1970's and early 1980's, the Naval Environmental

I. INTRODUCTION AND BACKGROUND

Numerical weather prediction models have made great progress in forecasting the basic meteorological variables and fields on the synoptic scale, such as sea-level pressure, wind, and moisture (e.g., relative humidity). However, dynamic models have had little success in predicting sensible weather variables at the regional/local scale, and in fact most models do not forecast many of these variables directly. Stochastic-dynamic prediction is being explored; it shows promise for operational use sometime in the future, but it awaits much further development and more powerful computers.

One of the most significant developments in weather prediction is the combination of dynamical and statistical methods, known as model output statistics (MOS). The MOS technique is the determination of a statistical relationship between a weather element of interest (e.g., visiblity, ceiling, precipitation) and a large menu of parameters output from an operational numerical prediction model (e.g., boundary layer wind, constant-pressure height, temperature). In the case of the National Weather Service, the operational MOS technique is based on multiple linear regression, where the prediction equations are developed from forecast model parameters (predictors) and observed weather (predictands)

7. PP--the potential predictability of the weather element by any given predictor. Potential predictability of a predictand/predictor pair is defined by Karl (1984) as

$$PP(2|1) = n/(n-1) \sum_{i=1}^{m} P_1(i)[\sum_{j=1}^{n} (P_{21}(j|i) - 1/n)^2]$$

where:

$P_1(i)$ = the marginal probability of a predictor;

$P_{21}(j|i)$ = the conditional probability of the jth predictand, given the ith predictor.

8. EPI--equally populous interval used to discretize the predictors (i.e., subintervals of equal population size based on the predictor range of values).

9. Functional dependence--a measure of the stochastic dependence of one predictor upon another. Functional dependence is the probability that one of the predictors will change when the other changes. High functional dependence values between one already selected predictor and another potential predictor indicates that little additional information beyond the first selected predictor is possible. Conversely, a low functional dependence value between the same two predictors, indicates that each predictor possesses distinct information about the predictand. Functional dependence range is 0.0 to 1.0 (1.0 = highest functional dependence). The specific derivation and mathematical description of the concept of "functional dependence" is discussed in greater depth by Preisendorfer (1983c).

10. Root-sum-squared functional dependence--the functional dependence of a predictor on all predictors already included in the developmental model. It is equal to the square-root of the sum of the squares of the individual functional dependence values.

11. TS1, TS2, TS3--threat score for weather element category I, II and III, respectively, computed from a contingency table (see Appendix E).

B. COMPUTER PROGRAMS

Four computer programs were developed by Karl (1983) to test the proposed Preisendorfer (1983 a,b,c) methodology for forecasting visibility. These programs were rewritten to allow them to be applied to cloud amount and ceiling forecasting and are on file in the Department of Meteorology, Naval Postgraduate School, Monterey, California, 93943.

1. A program to compute A0, A1, CE and PP for all predictors, all strategies (MAXPROB I, MAXPROB II and natural regression) for a particular number of equally populous predictor intervals. Statistics for the three strategies are based upon the predictor(s) that proved optimal for each strategy.

2. A program to compute functional dependence for all predictors, on a given predictor, for a given number of equally populous intervals and to compute the associated 96% critical confidence interval value (referred to as functional dependence(96) in this study) by Monte Carlo means.

3. A program to construct contingency tables and to compute skill and threat scores, for both the testing and training data.

4. A program to generate 100 random data sets, from the marginal probabilities of the predictor(s) in the developmental model, and to compute upper and lower 5% critical confidence interval values for A0 and A1 to be used for testing the significance of the results for each of the Preisendorfer models against chance. These confidence interval values are calculated via Monte Carlo means. This study developed another testing standard derived as a consequence of the central limit theorem. It is used in the results section to discuss the significance of the results of each of the models used, and is presented later in this chapter.

A second set of programs was used to develop the regression equations taken mainly from the BMDP STatistical Software Package (University of California, 1983).

1. BMDP P9R. An All Possible Subsets Regression program used to initially select predictors beginning with a general screening of the entire set of potential predictors.

2. BMDP P1R. A straight regression program to develop the prediction equation using the variables selected by the P9R.

3. BMDP P5D. This program takes the developed prediction equation and produces histograms of the data set divided into the prediction categories.

4. A program to generate the thresholds used with the regression equations. (These will be discussed in more detail later in this chapter.)

5. A program was developed to construct contingency tables of skill and threat scores from the regression equation experiments for both the training (dependent) and testing (independent) data sets.

C. MODELS

1. Preisendorfer PR Model

This model represents the first of two different applications of the basic Preisendorfer methodology (Preisendorfer, 1983 a,b,c). Karl (1984), in his preliminary research, provides a rigorous interpretation and results associated with this approach. Karl's study provides the necessary background for the continuing MOS studies using this model. This material will not be repeated here.

The PR model utilizes the working set of NOGAPS model output parameters (MOP's) and derived parameters (Appendix D) as potential predictors in constructing a developmental model, based upon the training data set, which provides the structure by which the testing data set is tested and evaluated. In general, these potential predictors have their range of values partitioned into discretized equally populous predictor intervals ("cells"), and conditional probabilities of

the predictand are calculated according to the three categories for cloud amount and ceiling, specified in Chapter III. Three separate strategies, for determining the specific category which is to be identified with each predictor value, are proposed. These strategies, two based upon maximum probability and the third based on a natural regression approach, are addressed as MAXPROB I, MAXPROB II and natural regression (NATR) in the remaining portions of the study.

Initial evaluation of this model involves varying the equally populous predictor intervals from sizes of four through ten, and selecting an optimal first predictor which provides one of the following requirements in the designated order:

a. the lowest CE value of all the potential predictors;

b. the highest PP value of all the potential predictors.

Once a first predictor is identified for each of the four through ten equally populous predictor intervals, corresponding category I, II and III threat and A0 skill scores (Appendix E) are calculated for both the dependent and independent data sets. The practice of selecting an optimal equally populous predictor interval (optimal in the sense of maximizing A0) from the eligible grouping sizes of four through ten, was proposed by Karl (1984) and used by Diunizio (1984) as a practical procedure which would permit the realization of peak skill scores as well as maintain associated computer storage requirements at a manageable level. An unfortunate consequence of this range of potential

grouping sizes is that certain statistical calculations associated with equally populous predictor intervals of eight, nine and ten are terminated before completion due to a two mega-byte storage ceiling at the NPS W.R. Church Computer Center (Diunizio, 1984). When considering potential predictor intervals, the size of the interval is of obvious importance, with lower values being the most desirable. In the previous studies in the MOS series concerning visibility, the criterion for determining the optimal equally populous predictor interval was to select the smallest interval value which maximized the dependent data set A0 and independent category I threat score. The threat score for category I was selected for this purpose because it was felt that low visibility (represented by category I) was uniquely important to forecast. In dealing with cloud amount there is not a single category that is obviously most important to forecase, and therefore, the selection of the interval was based only on the maximized dependent A0. This interval was then fixed for all ensuing aspects of the model evaluat[illegible] Consistent with the findings of the previous studies, t[illegible] selection criteria are based on the MAXPROB II scores, hence the MAXPROB I and natural regression strategies play no role in the predictor selection scheme.

Once the first predictor and its associated equally populous predictor interval have been identified, a functional dependent test of the first predictor against the remaining potential predictors is run. The second, third and all

subsequent predictors are selected only if both of the following criteria are met:

a. subsequent predictors must increase A0 over the A0 value attained at the preceding level, and

b. the selected predictor must have the lowest root-sum-square functional dependence of all the remaining potential predictors.

Significance tests were run on the developmental model after each predictor selection stage had been completed to determine if the results were suitably significant as compared to random chance. This was accomplished using the previously mentioned Monte Carlo method generating the 05 and 96 percentile confidence intervals using 100 randomly generated data sets. Further consideration has brought out that 100 cases may not be a large enough sample size for the Monte Carlo test. For this reason a testing technique, derived as a consequence of the central limit theorem more fully described at the end of this chapter, was applied to the results at the end of each run to demonstrate that the results are significant in relation to chance.

The model development continues along these criteria until computer storage limitations preclude further addition of parameters. This generally occurred in previous studies, and in every case in this study, at the fifth predictor level. Once the developmental model is completed, contingency tables of the forecast element category versus the observed element category are constructed for both the dependent and independent data sets, and threat and skill scores are computed and compared.

2. <u>Preisendorfer PR+BMD Model</u>

This model is still the PR model described above, but now sets of two linear regression equations are added to the list of potential predictors, namely, NOGAPS MOP's and derived parameters.

3. <u>Linear Regression Models</u>

Linear regression represents the more traditional approach to MOS. Regression Estimated Event Probability (REEP) is the basis for the National Weather Service and Air Weather Service regression models. In this study two approaches to the regression model are explored, a single stage and a two stage, and three threshold algorithms are used: equal-variance, quadratic, and a modified maximum-likelihood-decision-criteria. The procedures are outlined here, but a more detailed explanation of the theories is given in Appendix A.

a. Single Stage Regression

This model, referred to in the tables as BMD SS, consists of generating a single linear regression equation trained on the dependent data set, with the predictand set equal to 1, 2 or 3, corresponding to weather element categories I, II or III, respectively. This equation is then used with the dependent training set in the graphical plotting program BMD P5D, from the BMDP Statistical Software, to generate a set of three histograms and a listing of the individual frequency of observation (P), mean (u), and standard

deviation (s) of each of the three predictand distributions. These statistics are then used in the threshold algorithms to calculate two threshold values. Finally, the regression equation and the two thresholds are used to process the independent data to obtain a set of the observed weather element versus the forecasted element results in contingency table format. These tables and their calculated threat scores are presented in Chapter V and Appendix I.

b. Two-Stage Regression

This model, referred to in the tables as BMD TS, is based on a decision-tree scheme using two linear regression equations trained on the dependent data. The first equation is generated by separating the largest frequency category from the other two. In the cases of cloud amount and ceiling this was accomplished by setting the values for category I and II to 1 and the values for category III to 2 and then developing a regression equation and threshold (as in the single stage above) to suitably describe the two distributions. The second stage regression equation and threshold are generated, based only on those observations which did not exceed the first stage threshold value, effectively eliminating cases evaluated by the first stage as being category III. The second stage is thereby a separation of category I from category II observations. In other words, the first stage regression separates category III from the combined grouping of categories I and II, while the second

stage separates the remaining category II from category I data. The two resulting equations and associated thresholds are then applied to the independnet data to obtain the forecast versus observed contingency tables and calculate the threat scores.

c. Threshold Models

The equal variance model (referred to as EVAR) uses an algorithm which requires the assumption that the variances of the two normally distributed populations which are to be separated by a threshold are equal, while their means are unequal. The quadratic threshold (referred to as QUAD) algorithm makes no assumptions about the means and variances, but does take into consideration group apriori probability. The maximum likelihood decision criteria (MLDC) was modified for use as a third threshold model in order to separate the categories of scattered and broken clouds, categories I and II (Cooley, 1978), historically a difficult task. The MLDC is not based on apriori group probabilities but requires only the event conditional probability functions of the observations, and is useful in predicting events of rare occurrence. In the study, the MLDC threshold model consists of using the midpoint between the category I and II distribution means with the EVAR threshold between the category II and III distributions.

D. SIGNIFICANCE TESTING

The results of the experiments are tested against two standards to demonstrate that the results are significant

with respect to chance and to evaluate improvement over classical MOS modeling methods.

1. Significance of the Skill of a Forecast versus Chance

This first test of the results is based on the proposal that both percentage correct (A0) and threat scores (TS1, TS2, TS3) can be presented as probabilities and the fact that a binomial population, if large enough, can be approximated by a normal distribution. As such the percentage correct and threat scores may be subjected to a null hypothesis significance test derived as a consequence of the central limit theorem. The actual significance testing is made with respect to confidence intervals about the scores which would be achieved by a uniform random distribution of category I, II and III observations in a 3×3 contingency table. These scores represent the scores which would be achieved by pure chance. The test can be stated that the null hypothesis is

$$P(A \leq X \leq B) = .95$$

with lower limit

$$A = y/n - 1.96[y/n(1-y/n)/n]$$

and the upper limit

$$B = y/n + 1.96[y/n(1-y/n)/n] .$$

A 95% confidence interval is made about the scores expected if each category had an equally likely chance of being forecasted. If the procedure score lies within the interval for the null hypothesis score then it is considered that there is no statistically significant difference between the two scores. The contingency tables and 95% confidence interval calculations are shown in Figs. 3, 13, 20, 27 and 37.

2. Improvement Over Baseline

Since some form of regression is the traditional method of developing MOS models, the baseline standard for comparison of all the experiments in this study are confidence intervals generated using the results from the single-stage regression for each area, and time period. These 95% confidence intervals are made using the same equations as the test for significance of a score versus chance. Additionally, in area 2, the TAU-00 baseline is used to evaluate degradation of the results with time. The baseline intervals are shown in Figs. 10, 17, 24, 30 and 37.

E. MEASURES OF SEPARABILITY

As the testing proceeded through progressive time stages, it became more apparent that the methods were struggling to separate the categories of scattered and broken clouds, categories I and II (Cooley, 1978). This problem required investigation of some alternate predictor selection schemes to improve the ability to discriminate between these categories. Two approaches of determining the optimal separation

between the categories were combined and then applied in a brief analysis on area 2, TAU-00. These methods are termed Class Separability Measures and Cluster Analysis. Unfortunately, time did not permit a detailed attempt at using these two methods, but the results from area 2, TAU-00 are included in this study and show sufficient potential to deserve further study.

1. <u>Class Separability Measures</u>

The specific separability measures used were the Bhattacharya Distance, the Divergence, and the Mahalanobis distance (Hand, 1981), each of which is discussed in more detail in Appendix B. These measures were calculated using the means and variances, and in the case of the Mahalanobis, pooled variances of the various predictors with the following univariate form:

Bhattacharya Distance:

$$Bh = \frac{(\mu_1 - \mu_2)^2}{}\left(\frac{2}{\sigma_1^2 + \sigma_2^2}\right) + \frac{1}{2} \ln \left(\frac{\sigma_1^2 + \sigma_2^2}{2\,\sigma_1 \sigma_2}\right)$$

Divergence:

$$Div = \frac{2(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} + \frac{1}{2}\left(\frac{\sigma_2^2}{\sigma_1^2} + \frac{\sigma_1^2}{\sigma_2^2} - 2\right)$$

Mahalanobis:

$$Mal = \frac{(\mu_1 - \mu_2)^2}{\sigma^2}$$

The predictors were separated by classes. In this case categories I, II and III, and then the three measures were calculated for category I versus II, category II versus III, and category I versus III. It is important to note that the variables that best discriminate between group I and II may not be the same as those that best discriminate between II and III or between I and III. In each case the means and variances of the predictors were scaled from 0 to 100 to ease number handling and value comparisons. The calculated distance measures are listed in Tables X, XI and XII for area 2 at TAU-00.

2. Cluster Analysis

Cluster Analysis takes a sample of potential predictor variables of unknown classification and groups those variables into natural classes or "clusters." The method is fundamentally a tool for data exploration to determine if natural and useful groupings do, in fact, exist. This method was applied to the predictors by use of the BMDP Statistical Program, P1M, which provides four measures of similarity for clustering variables and three criteria for linking or combining clusters. A more detailed discussion of cluster analysis is also found in Appendix C. In general, clustering was used to determine groupings of predictors that carry much the same information in relation to the predictand classes.

3. <u>Experiments Using Separability and Clustering</u>

It must be noted that time did not permit an extensive investigation of these methods, but rather only a cursory look at their potential for usefulness. The basic method consists of using the cluster analysis to develop groups of variables to choose from, and then employs the separation measures to select the "best" predictor from each of these clusters. These parameters are then used to develop a linear regression equation to predict the three cloud amount categories in the same manner as earlier testing in this study. Four experiments were attempted:

a. A single-stage regression using the variables which had relatively high separability measure values for category I versus II.

b. A single-stage regression using the variables which had relatively high separability measure values for category II versus III.

c. A single-stage regression using the predictors with the highest separation value from each clustered group of predictors.

d. A two-stage regression using separation and clustering to separate category I from II and III and then category II from III.

[illegible]. GENERAL

The first area studied was cloud amount in area 4, TAU-[illegible], and the procedure is an exact application of the methodology used in the previous MOS studies for visibility (Karl, 1984; Diunizio, 1984). The one exception is that the linear regression model is tested with both a single-stage and two-stage regression technique. Next, area 2 of the

its TS2 falls below the baseline significance. Its maximum A0 of 49.49 is not significantly different than baseline and is attained at the first predictor level. By the fourth predictor it has reached a TS1 that is significant both to chance and the baseline but only with severe degradation to both its A0 and TS3 scores. NATR does very poorly overall, attaining its maximum A0 (45.62) at four predictors, which is within baseline interval, but only marginally within the TAU-00 baseline confidence interval. Unlike MAXPROB I and MAXPROB II, NATR is not able to predict category I with any acceptable credibility, even after four predictors. It, too, retained TS2 and TS3 values that are not significantly different than the TAU-00 baseline.

On the other hand, the PR+BMD model (Fig. 26 a-c) produced very different results. It selected a grouping size of six equally populous intervals, reaching its peak A0 for MAXPROB I and II at the second predictor. While slightly lower than the A0 for PR, the identical results of MAXPROB I and II show some skill at forecasting category I. MAXPROB I shows a TS1 of .13 and MAXPROB II shows a .17, both of which are significant improvements over the baseline, and in the case of MAXPROB II is marginally significant with respect to chance. Although still lagging behind in A0 by nearly 3%, NATR also shows significant improvement over baseline in TS1 but not enough to be considered significant with respect to chance.

Baseline Interval

A0 : 42.97 to 50.97

TS1: .00 to .02

TS2: .32 to .39

TS3: .36 to .44

The MLDC moves the threshold to 1.91, resulting in a decrease in A0 of nearly 5%, falling outside the baseline confidence interval. Although the TS1 was raised to .17, it is not significant with respect to chance, and the TS2 suffered severe degradation such that it is no longer significant with respect to chance either.

The PR model (Fig. 25 a-c) selected eight for a grouping size, which limited the model to only four predictors due to a 2 megabyte limitation at the NPS computer center (this is addressed in Chapter IV of this paper and in Diunizio, 1984). All three strategies in this time period suffer the same inability to forecast category I cloud amount. It is not until the fourth predictor that any of the scemes, namely NATR attains higher than a .04 TS1. MAXPROB I attains its relatively high A0 peak (50.31) at the second predictor, but is unable to forecast any category I at this level. By the fourth predictor it attains a TS1 of .13 while droppings its A0 to 47.25 and its TS2 (.27) below baseline significance. MAXPROB II strongly overpredicts the category III overcast situation, which gives it a TS3 value of .45. While this is a statistically significant improvement over the baseline,

The BMD single-stage regression (Figs. 21-23) starts to show signs of deterioration with time as one would expect. Again, EVAR has the highest A0 at 46.97, which is significant with respect to chance and is not significantly different from the TAU-00 baseline, but it is near the lower limit of that baseline confidence interval. Most of the degradation takes place in the TS1 category, which is not doing well at TAU-00, but is doing even worse at TAU-24. BMD EVAR and QUAD are almost unable to distinguish any category I observations from category II (TS1 of .01). At the same time both TS2 and TS3 remain within the confidence interval for the TAU-00 baseline, showing no significant difference. The BMD equation yields an even smaller separation between the means of category I and II than was seen in TAU-00. In this case, the mean for clear scattered case is 2.068 with a standard deviation of .232, and for the broken group, the mean is .214 with a standard deviation of .223. The obvious problem here is that the separation between the means is less than one-third that of the standard deviations! This is a tough problem for any threshold model. EVAR and QUAD produced thresholds of 1.679 and 1.614 respectively, both well left of the mean of the scattered cloud group, accounting for the almost zero forecasting of category I cloud amount. The BMD EVAR results lead to the following baseline intervals (Fig. 24):

NATR on the other hand attained the highest A0 yet received by any method or area at 53.43%. This fell .01% above the baseline interval and therefore could be considered to be a marginally significant improvement over the baseline BMD model. NATR also showed significant improvement over baseline in TS1 (compare .42 with the upper limit of .39). Although NATR remained within the interval of significance for the baseline in TS1, it still fell short of statistical significance with respect to chance.

It is clear by all measures that the PR+BMD method, specifically the MAXPROB I strategy, achieved the best results. It is significant that none of the methods could forecast category I cloud amounts with a skill level better than pure chance.

2. Area 2, TAU-24 (Table VII)

The TAU-24 time period has an extra five Model Output Parameters (MOP's) added to the available predictors. All other MOP's and derived parameters remained the same (see Appendix D).

The following is the confidence intervals for significance with respect to chance (Fig. 20):

Significance Test

A0 :	29.43 to 37.35
TS1:	.11 to .17
TS2:	.18 to .25
TS3:	.20 to .27

but also attaining better scores than baseline in all four scores. This improvement over baseline however is not enough to be above the confidence interval for improvement, although nearly so in both TS1 and TS2. NATR did not attain peak A0 until the third predictor (48.07) but still 3% lower than MAXPROB I or II. NATR displayed an actual significant improvement over the baseline interval in the TS2 score (.42) but fell below baseline significance in TS3. When allowed to progress to five predictors not only did NATR continue to improve at the next step, but also each of the three schemes improved in TS1 (MAXPROB II scored a TS1 = .20 at the fourth predictor) while degrading TS2, TS3 and A0. This might be significant at some time if TS1 were decided to be the most important category to forecast.

The PR+BMD, Fig. 19 a-c, selected a grouping size of six, and attained peak A0 for MAXPROB I at four predictors, and for MAXPROB II and NATR at three predictors. MAXPROB I attained the same A0 as it did in the PR model but improves its TS1 and TS3 scores. Although TS1 did not improve to significance with respect to chance, it did improve significantly over the baseline (compare .17 to .11). TS2 was nearly equal to baseline, but TS3 was at the upper limit of the baseline confidence interval for improvement. MAXPROB II did not fare as well as MAXPROB I overall but TS2 and TS3 did remain within the confidence interval of the baseline, showing no significant difference or improvement.

has a mean of 2.139 with a standard deviation of .246. With the means of the two categories being separated by less than one-half of a standard deviation, it is easy to see why the TS1 is so low. Both the EVAR and QUAD models place the threshold value separating category I from II well to the left of the mean of category I (EVAR threshold = 1.705 and QUAD threshold = 1.643). This situation holds throughout the area 2 testing of the BMD model. The QUAD model shows only slight variation from EVAR, which is not surprising since the thresholds are very nearly the same. The MLDC model moves the threshold between category I and II to 1.864. This significantly raises the TS1 above the testing confidence interval, but loses 2% on A0 and nearly reduces TS2 below significance levels. The resulting baseline confidence intervals from Fig. 17 are:

Baseline Interval	
A0 :	45.40 to 53.42
TS1:	.07 to .11
TS2:	.32 to .39
TS3:	.36 to .44

The PR model, Figs. 18 a-c, selected a grouping size of six equally populous intervals and achieved its peak A0 at the second predictor level with MAXPROB I and MAXPROB II (51.26). Both schemes have the same results at this stage, not only showing significant results in A0, TS2 and TS3

improvements accomplished by the other methods in that time period. The TAU-00 baseline is used to compare all three time frames so that a trend with time can be evaluated as well.

1. Area 2, TAU-00 (Table VI)

The significance test with respect to chance is calculated in Fig. 13 and yields the following intervals:

Significance Intervals

A0 : 29.52 to 37.08

TS1: .12 to .18

TS2: .18 to .25

TS3: .19 to .26

The first model evaluated is the BMD single-stage regression using the EVAR threshold, the baseline for evaluating other models (Fig. 14). It produced an A0 that is significantly better than chance (49.41% compared to 37.08%) and very significant values for TS2 and TS3. In fact, these TS values exceed the EVAR model in area 4, TAU-00. However, the price is paid in the TS1 value. The EVAR model was able to obtain only a .09 threat score for the clear/scattered category, obviously well below the significance test for chance. This is the result of the BMD equations being unable to clearly separate the clear/scattered category from the broken category. The historgrams show that the equations result in category I having a mean of 2.033 and a standard deviation of .257, while category II

significantly different (although lower) than the baseline interval. Both MAXPROB I and II improved over their PR counterparts in every score except TS3.

An interesting difference between the Preisendorfer models versus the linear regression models that holds throughout the study is in the response of the dependent scores. In the BMD models the dependent data (training set) scores are very near to those of the testing (independent) scores, whereas in the Preisendorfer schemes the dependent A0 scores typically rise to values above 90% with the addition of the fifth predictor. This may indicate that the PR models do an excellent job of fitting the training sample but do not make proper inference concerning the structure of the population from which the sample was drawn.

B. NORTH ATLANTIC OCEAN AREA 2 CLOUD AMOUNT

Area 2 (Fig. 2) encompasses a geographic region that extends from the southeastern tip of Newfoundland, across the North Atlantic Ocean to the eastern coast of England, north to the Five Fingers of Iceland and back to the Canadian coast north of Newfoundland. Area 2 was studied through all three time periods, TAU-00, TAU-24 and TAU-48, and each will be discussed separately. As in area 4, a null hypothesis is generated for each time period to evaluate the significance of the results versus chance. Also, as in area 4, a set of confidence intervals based on the BMD SS model for each time period is used as the baseline for measuring

The PR model results are shown in Figs. 11-12. The model selected a grouping size of six and the peak results were attained by all three strategies at the five predictor level. Natural regression (NATR) produced the highest A0 (45.36) for this model, followed by MAXPROB I (44.14) and MAXPROB II (41.14). Though all of the strategies had significant A0 values compared to chance, none of them improved on the A0 of the baseline. MAXPROB I improved on BMD for TS3, but lagged in other scores. The scores all lie in or below the confidence interval for the baseline and, therefore, cannot be considered to be significantly different. MAXPROB II did appreciably worse in that it showed significance with respect to chance but its A0 and TS2 were below the baseline interval. NATR, with the best A0 of the three PR strategies, lost skill in category I, as indicated by TS1, and this is not even significant with respect to chance. Its TS2 and TS3 were not significantly different than the baseline values. It is of interest to note that the PR scheme and the BMD single-stage model did not select any common predictors.

The PR+BMD scheme selected a grouping size of six and attained peak A0 values for MAXPROB I and II at two variables and for NATR at three variables. In this case both the MAXPROB I and II produce near equal results, with MAXPROB II showing slightly higher A0, TS1 and TS2. In all three cases the A0 was significant compared to chance but not

again, by moving the thresholds with the MLDC model, a decrease in A0 is observed (43.67%) with significant increases in TS1 and TS3. This increase in TS1 and TS3 is also at the expense of decreasing TS2 below significance levels.

Because the results of the single-stage regression were so much better than the two-stage, and in view of the fact that all the homogeneous areas of the North Atlantic Ocean dispaly similar distributions, the single-stage model was pursued for the remainder of the cloud amount experiments. A single exception will be discussed later. In Chapter IV it was mentioned that because linear regression, of some form, is the traditional method for MOS studies and operational models, it would be selected as a baseline measurement (in addition to the confidence interval generated by the null hypothesis contingency table) to measure the skill of the other methods. The single-stage BMD with the EVAR threshold model was selected as this "baseline" measure. Fig. 10 shows the development of the confidence intervals for area 4 baseline. The resulting intervals are:

Baseline Intervals

A0 : 43.21 to 49.19

TS1: .25 to .31

TS2: .30 to .36

TS3: .24 to .30

However, the threat scores for category I and III are below the confidence interval for pure chance. Th poor results are most likely a reflection of the particular nature of the frequency of occurrence in each of the observation categories. The two-stage regression was chosen in the visibility studies because of a very low occurrence (most cases less than 5% of total observations) of low visiblity. Since low visiblity was the threat most desired to predict, the two-stage regression was chosen to more skillfully predict a low frequency category. In area 4, on the other hand, the frequency of observation is nearly the same for all three categories of cloud amount. The thresholds of the two stages were moved closer to the middle in the MLDC model (Fig. 6) in order to better predict the outside two categories, I and III. The resulting A0 is 2% lower than the EVAR or QUAD models, and an increase in threat scores for both categories I and III occurred. Only the new threat score for category I (.33) increased beyond the significance level, but a large price was paid in the TS2, which dropped to .27, close to the significance-level boundary.

Since the frequencies of occurrence for the three categories are nearly equal, a single-stage regression model was next attempted (Figs. 7-9). The EVAR threshold model demonstrates only a 1.0% increase in A0, but much more importantly, all three categories have threat scores significantly above chance. The QUAD model has a slightly higher A0 (46.67%) than EVAR and very similar threat scores. Once

attained by that particular strategy, while the dependent scores reflect the results attained using the first five predictors selected.

A. NORTH ATLANTIC OCEAN AREA 4 CLOUD AMOUNT

Area 4 was selected as the first for evaluation because of its large sample size and nearly equally populous observation categories I, II and III. This area encompasses a broad region of the North Atlantic Ocean with the southern border reaching to the northeastern tip of Portugal and extending northward through the English Channel to encompass the southern portion of the North Sea (Fig. 2).

1. Area 4, TAU-00 (Table V)

The following are the confidence intervals for significance with respect to chance (Fig. 3):

Significance Intervals	
A0 :	30.53 to 36.19
TS1:	.17 to .22
TS2:	.20 to .25
TS3:	.15 to .20

The first model tested is the two-stage BMD in the same manner as the previous NPS visibility studies (Karl, 1984; Diunizio, 1984). The results, s. n in Figs. 4 and 5 (EVAR and QUAD), show an A0 of 45.27% which is significant with respect to chance and category II threat score (TS2) of .40 which is also highly significant compared to chance.

V. RESULTS

The procedures for the experimentations on predicting cloud amount and ceiling, as specified in Chapter IV, were followed for the North Atlantic Ocean homogeneous areas 2 and 4. These homogeneous areas are displayed in Fig. 2. The results of these procedures are summarized in Tables V through IX, and detailed results are displayed in Figs. 3 to 44. This chapter discusses the results and significance of each area and each model run using the information on these figures. Cloud amount is pursued first in the study since it is important to the prediction of ceilings, as noted in Chapter III.

The terms used throughout this section are defined in Chapter IV. The linear regression models are referred to as BMD and the three threshold models are Equal Variance (EVAR), Quadratic (QUAD) and Maximum Likelihood Decision criteria (MLDC). The Preisendorfer method is used both with (PR+BMD) and without (PR) linear regression equation predictors. In each model A0 (total percent correct) is used as the criterion for the "best" model. In the PR and PR+BMD models, a contingency table is generated for all three strategies, MAXPROB I, MAXPROB II and natural regression, with the addition of each new predictor. In all cases, the independent score discussed reflects the best score

North Atlantic Ocean was tested at all three time increments, TAU-00, TAU-24 and TAU-48, in the same manner, but without the two-stage regression. Finally area 2, TAU-00, was tested using the measures of separability and clustering techniques.

Testing on ceiling height prediction was limited to area 2, TAU-00, using initially the same methodology. An experiment was then made to test the ability to forecast ceilings given perfect skill at predicting cloud amounts. In this case the categorized cloud amount was used as a predictor in the ceiling prediction methodologies. The results of each of these tests are discussed in the next chapter, and are summarized in Tables V through IX and Figs. 4 through 44.

In general, it can be observed that an expected degradation is experienced from TAU-00 to TAU-24 in all of the methodologies. The inability to forecast category I is the most glaring problem. At TAU-00 the skill levels are poor in forecasting category I, but in TAU-24 they become nearly zero in all but the PR+BMD method. In no case are any of the methods able to attain significant skill in forecasting scattered clouds in comparison to pure chance. However, it would be fair to observe that the A0 degradations in general are not as significant as one might have expected.

3. Area 2, TAU-48 (Table VIII)

The area 2 TAU-48 time period also has the five extra predictors mentioned above in TAU-24. The following is the confidence intervals for significance of the skill scores with respect to chance for TAU-48, area 2 (see Fig. 27):

Significance Test

A0 : 29.30 to 37.11

TS1: .12 to .18

TS2: .18 to .25

TS3: .20 to .26

As one would expect, the models continue to experience a degradation with time. The BMD EVAR model attains only an A0 of 45.32 which, although well above the significance test for chance, still falls below the TAU-00 baseline confidence level indicating that it is significantly worse, and

that considerable degradation has occurred. Additionally, the threat scores for category II and III are only marginally within the TAU-00 baseline confidence. Interestingly, though only by a small amount (TS1 = .04), the BMD TAU-48 is able to forecast category I better than TAU-24. The overall degradation in performance is clearly seen in distributions of the three categories by the BMD equation. Between the category means for I and II there is now only a separation of .05 while the standard deviations are of the order of .18. This same degradation is seen in the separation of categories II and III, where the means are now 2.192 and 2.270 and the standard deviations are .178 and .189, respectively. This shrinking of the separation of the means is to the point at TAU-48 that the QUAD model is unable to produce a non-imaginary threshold between category I and II. MLDC also performs consistent with previous time periods, this time reducing the A0 to 42.93 which is only 6% better than the upper limit on chance. In fact, only the MLDC TS3 proves to be significantly better than the pure chance contingency table. These results lead to the following TAU-48 baseline interval (Fig. 30):

<u>Baseline Interval</u>

A0 :	41.31 to 49.22
TS1:	.02 to .06
TS2:	.28 to .36
TS3:	.32 to .40

The PR model (Fig. 31 a-c) also experiences the same degradation with time. The model selected grouping size seven and reached peak A0 at three predictors for MAXPROB I, two predictors for MAXPROB II and four predictors for NATR. MAXPROB I loses 3.5% in A0 from TAU-24, and also drops below the significance limit for baseline TS2. At the same time though, it improves on the baseline for TS1, though not enough to be considered significant with respect to chance. MAXPROB II fares somewhat worse in every category except TS2 where it maintains a score within the baseline interval. When compared for time degradation with the TAU-00 baseline, it is only marginally within the baseline interval for A0 and TS1 and just below for TS3. Likewise, NATR scores are within the 48-h baseline interval, with the exception of TS1, but are significantly worse than the TAU-00 baseline in every category with the exception of TS1.

The PR+BMD selected a grouping size of six and reached its peak A0 at the first predictor level. The identical scores of MAXPROB I and MAXPROB II show statistically significant improvement over the TAU-48 baseline in both A0 and TS3. However, the TS1 scores of zero reveal its inability at this time period to forecast category I. When the model runs out to five predictors, where NATR peaks on A0, then it can be seen that all three schemes forecast category I with a TS1 equal to or in excesss of .20. For example, MAXPROB II at the fourth predictor has an A0 of 45.94, but

is significant in all threat scores with respect to chance (i.e., TS1 = .26, TS2 = .28, TS3 = .34). NATR again performs well below the MAXPROB strategies, even though its A0, TS2 and TS3 scores are within the baseline confidence interval.

In general, it can be said that all the schemes suffered significant degradations due to the 48-hour time period. Forecasting category I (scattered/clear) remains a problem through all time periods, and is only forecastable at large cost to the other threat scores and the total percentage correct.

4. Area 2, TAU-00 Experiments in Clustering and Separability (Table VII)

A brief description of the separability and cluster methods and procedures is found in Chapter IV, and a more detailed theoretical description is found in Appendices B and C. The results of the measures of separability program are listed in Tables X-XII and the clustering of variables is listed in Appendix C. The baseline for comparison in these examples is the area 2, TAU-00 BMD using the EVAR threshold, and the null hypothesis significance confidence intervals used for TAU-00, area, 2.

The first test consisted of selecting the predictors from the category I versus II grouping of the measures of separability, using those predictors with the highest divergences. As Table X shows, the values for the three measures were very low in this grouping, which is a possible clue to the low skill attained by all the methods in category

I. Predictors were chosen that had a divergence of .08 or higher; predictors were then used in the BMD EVAR model. The results of the test are shown in Fig. 33. The model attained an A0 of 45.39 which is significant with respect to chance, but is outside the low end of the confidence interval for the baseline. The TS2 is not significantly different from baseline but TS3 (.34) fell just below the lower limit of the baseline value. Most importantly, though, the model did not predict any category I. Although this is well below the baseline value, the baseline values are significantly worse than chance.

The second test, found in Fig. 34, is similar to the first, with the exception that the variables were selected from the category II versus III grouping of the measures of separability (i.e., predictors with values above .35). These measures showed much higher values, which is consistent with the results of the methods in area 2, TAU-00 where the threat scores for category II and II are very much higher (i.e., the models are able to separate II from III much easier). This time the A0 improved to 48.91, nearly equaling the baseline value. The model also equalled baseline performance in threat scores TS2 and TS3. Once again, however, the model is unable to forecast any category I observations.

The third test tries to combine the clustering information with the measures of separability. In this case the clusters, listed in Appendix C, were used as the initial sorting of predictors. Next, the predictor from each cluster

that has the highest divergence in the category I versus III grouping, was selected as a variable. These variables were then used in the BMD EVAR and the results are shown in Fig. 35. This model did rather poorly and only stayed in the confidence interval for the baseline in TS2. All other scores dropped below the intervals for baseline while remaining significant with respect to chance (except for TS1).

The fourth test attempted to utilize all the information available. The clustering technique was combined with the measures of separability to select variables that would best separate category I from II, and then those that would best separate category II from III. These two sets of predictors were then used in a two-stage regression first separating category I from II+III and then II from III. The results, shown in Fig. 36, show much improvement over the previous three tests. In fact, this model produced the highest A0 attained by any of the BMD models so far studied. The EVAR threshold produced a 50.59 A0 which is higher than baseline but not significantly so, and TS2 showed modest improvement over the baseline interval. This model also produced a smaller TS1 (.05) than hoped for, but the fact that it is greater than zero is encouraging.

It is unfortunate that more time was not available to pursue further these methods, but the initial testing shows some potential for usefulness in the MOS methods. There are several important points to be made. First, the results of

the measures of separability program confirms that the regression methods used in area 2 are not forecasting category I with much skill, because the available predictors do not have enough information. Secondly, the clustering proves to be more valuable if the predictors are scaled some way to prevent all the velocity predictors being clustered and the height predictors being clustered, etc. (It is possible that this type of result is not due to scaling but rather to characteristics of the model producing the parameters). Thirdly, the measures of separability give high values to most of the predictors chosen by the two methods generally used in this study. That lends plausiblity to its usefulness as a predictor screening agent to reduce the number of predictors being forced through the various prediction strategies.

C. NORTH ATLANTIC OCEAN AREA 2 CEILINGS

The first experiments in forecasting ceiling were carried out using a direct application of the methods employed for cloud amount and previously for visibility. The frequencies of distribution of ceilng observations for the North Atlantic Ocean are shown on Table IV. Area 2 was chosen for experimentation, consistent with the concentration of MOS visibility and cloud amount effort. The second set of experiments is designed to evaluate the skill of forecasting ceiling given that there exists perfect skill at forecasting cloud amount. The cloud amount observations are then categorized and used as a parameter in the various methods. As in the studies on

cloud amount, a null hypothesis is established, using a contingency table, based on each category having an equal probability of being forecasted for each observation. This yields the following 95% confidence intervals for evaluating the significance of the results (see Fig. 37):

<u>Significance Intervals</u>

A0 : 29.23 ro 36.77

TS1: .13 to .19

TS2: .20 to .27

TS3: .17 to .23

1. <u>Area 2, TAU-00 Ceiling Tests Without Cloud Amount (Table IX)</u>

The results of the BMD single-stage regression is shown in Fig. 38. The resulting A0 is significant with respect to chance and is very similar to the values obtained in the cloud amount studies. Threat scores for category I and II are both well above significance with TS2 being the highest at 0.42. However, the single-stage model is unable to discriminate between category II and III. This is shown in the TS3 score of .00 and in the histograms displaying the distributions of the BMD equations. The means have good separation between category I and II (1.666 and 1.861 respectively, with standard deviations of .254 and .225). The problem occurs between category II and III where the means and standard deviations are 1.852 and 0.192 for category II and 1.873 and 0.183 for category III. The mean separation

is nearly one-tenth that of the standard deviations. The separation is so small that the QUAD model is unable to resolve a non-imaginary threshold. In this case the MLDC model shows promise. By moving the threshold halfway between II and III, TS3 increased from .00 to .15, which is not enough to be significant compared to chance, but it is noteworthy that the A0 also increased by 1% and there is little effect (-.03) on TS2. As in the cloud amount studies, the BMD single-stage regression will be used as the baseline for evaluating other methods. In this case, however, the BMD with MLDC threshold will be used. This produces the following confidence intervals (see Fig. 40):

<u>Baseline Interval</u>

A0 : 41.56 to 49.56

TS1: .19 to .25

TS2: .35 to .42

TS3: .12 to .18

The PR+BMD model chose a grouping size of six and attains peak A0 for the MAXPROB strategies at the second predictor and for NATR at the third predictor (Fig. 41 a-c). MAXPROB I achieves the highest A0 (47.24) of the three strategies and shows very different results in the threat scores compared to the baseline. It scores well above the baseline interval for TS1 and TS3, while showing only marginal improvement over chance in TS2. The most striking fact is that the PR+BMD does so well in the category III

(TS3 = .36) compared to the counterpart scores in the BMD model. MAXPROB II shows similar results, although slightly better in TS1 but slightly poorer in A0, TS2, and TS3. Actually, MAXPROB II shows no skill compared to chance in category II. For NATR, A0 peaks at the third predictor and attains the second highest percentage correct. In general, it does much poorer than the MAXPROB strategies, giving results that more closely resemble the baseline results. NATR shows no significant difference from baseline in either A0 or TS2, while scoring significantly higher in TS3 (though only on the margin of being significant with respect to chance). The TS1 of 0.18 is not significant compared to chance, but it is significantly worse than baseline.

In general, the methods applied to ceilng heights produced very similar results to those attained for cloud amount, both in percentage correct and in threat scores. BMD with MLDC or even EVAR does the best in forecasting category II but is poor in forecasting category I or III. Conversely, the MAXPROB strategies are much better at forecasting categories I and III, but at a cost of reducing the results for category II below significant levels.

2. <u>Area 2, TAU-00 Ceiling Using Cloud Amount Observations</u>

In these experiments cloud amount observations were categorized and used as a predictor for ceiling.

The results of the BMD model using cloud amount (Figs. 5, 42-43) are excellent compared to the results attained so far in this study. The EVAR model attained an

A0 of 67.50 which is 18% higher than the baseline BMD. The threat scores are consistently high as well, far exceeding the baseline in every category by .10 to .44. QUAD provides nearly identical results across all categories.

The PR+BMD model is then run making cloud amount available as a predictor (Figs. 5, 44 a-c). The model chose grouping size six and attains peak A0 at the second predictor. Cloud amount is the first predictor chosen and the linear regression equation variable (not containing cloud amount) is the second. MAXPROB I and MAXPROB II produce identical results at this level with an A0 of 68.60, about 1.0% higher than the BMD model using cloud amount. The model's TS3 is an outstanding .72 and TS2 is .49, both of which are very significant with respect to chance and the baseline. However, the resulting TS1 is only .24 which is significant compared to chance but is not an improvement over the baseline. The NATR model attains peak A0 at the fourth predictor, and does not perform as well as MAXPROB in any category. The TS3 of .64 and the TS2 of .451 show significance with respect to both chance and the baseline, but the TS1 value is not significantly different than baseline, and only marginally significant compared to chance.

The value of good skill in predicting cloud amount in forecasting ceiling is obvious by the above results. The very high category III results (i.e., ceiling greater than 3500 feet or unlimited) is probably due to the definition

of ceiling, namely that if cloud amount is less than 5/8 then the ceiling is unlimited. The strongest effect of cloud amount in the forecast of ceiling is whether or not a ceiling exists, thus the high threat score of category III which contains all the observations of no ceiling.

VI. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

The primary objective of this study was to begin the investigation into statistical forecasting of cloud amount and ceiling by extending the methods researched by Karl (1984) and applied by Diunizio (1984) in the area of visibility. The ultimate goal is to develop a viable statistical forecasting scheme suitable for eventual employment in an operational U.S. Navy marine ceiling and cloud amount MOS forecasting system. This is certainly not an exhaustive study of the subject, but does provide an important first step in statistically forecasting these weather elements.

The results of the tests in the various areas and time periods show that the methods evaluated are useful in forecasting both cloud amount and ceilings. Although the models are not yet producing results as good as one might desire for an operational MOS system, they are forecasting significantly better than pure chance, giving them useful skill levels. In area 4, TAU-00, the single-stage linear regression performed the best, and became the "baseline" from which to measure the other methods. In area 2 the model that scored consistently highest in all three time periods is the PR+BMD. The general problem experienced by all the approaches in area 2 is the inability to forecast the scattered/clear condition (category I) with any skill. Significant skill in this category is only attained in a very few cases and then only

: a great cost to the threat scores of the other two cate-
)ries. In the initial ceiling studies, PR+BMD gives the
est overall results, but the linear regression is able to
ore skillfully forecast the category II. In contrast, when
perfect cloud amount forecast is added as a predictor to
ie ceiling models, linear regression gives much better
esults overall, especially in forecasting low ceilings.

In the previous MOS studies in this series, a low visi-
ility situation was clearly the most threatening category
or operational Naval forces and, therefore, was selected
s the criterion to maximize as well as to evaluate one model
gainst another. In cloud amount predictions, there does
ot exist a single category that is clearly more important
han the other two. In the absence of a better measure,
bsolute percentage correct was utilized. The study does
eveal a need to develop some evaluation criteria for contin-
ency table output for the MOS project in general. This
ould be of great assistance in the developmental stages of
arameter selection as well as evaluating the overall per-
ormance of a particular model. The two measures used in
his study to evaluate significance of the results proved
o be very useful. The previous studies based significance
esting on a Monte Carlo scheme evaluating a set of 100
andomly generated data sets to produce upper and lower .05
ritical values for A0. The significance test used in this
tudy, derived as a consequence of the central limit theorem,

$$\Lambda(z) \;=\; \exp\{1/2 \left[\left(\frac{z-\mu_0}{\sigma_0}\right)^2 - \left(\frac{z-\mu_1}{\sigma_1}\right)^2 \right]\} \underset{c=0}{\overset{c=1}{\gtrless}} \frac{p_0 \sigma_1}{p_1 \sigma_0}$$

re $k = (2\pi)^{-1/2}$. Algebraic manipulation produces

$$(\sigma_1^2 - \sigma_0^2) z^2 + 2(\sigma_0^2 \mu_1 - \sigma_1^2 \mu_0) z$$

$$+ [(\sigma_1^2 \mu_0^2 - \sigma_0^2 \mu_1^2) - 2\sigma_0^2 \sigma_1^2 \ln(p_0 \sigma_1 / p_1 \sigma_0 0] \underset{c=0}{\overset{c=1}{\gtrless}}$$

ch is recognizable as a quadratic equation in z.

$$z^* \;=\; -b \pm (b^2 - 4ac)^{1/2}/2a$$

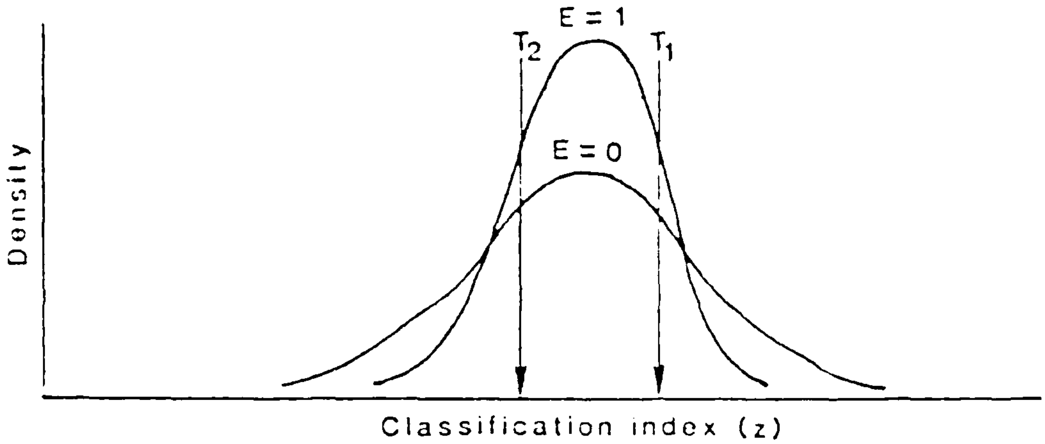
re:

$$a \;=\; \sigma_1^2 - \sigma_0^2$$

$$b \;=\; 2(\sigma_0^2 \mu_1 - \sigma_1^2 \mu_0)$$

$$c \;=\; (\sigma_1^2 \mu_0^2 - \sigma_0^2 \mu_1^2) - 2\sigma_1^2 \mu_0^2 \ln(p_0 \sigma_1 / p_1 \sigma_0)$$

Note that in this situation there are two thresholds. The group having the smaller variance will lie between the two thresholds.



The thresholds shown are typical of a situation where $p_1 < p_0$. Note that these thresholds lie between the two intersections of the densities. If the inequality of prior probabilities were reversed, the thresholds would lie outside of the region between the two density intersections. Further, note that the decision region for the group having the lesser variance lies between the thresholds.

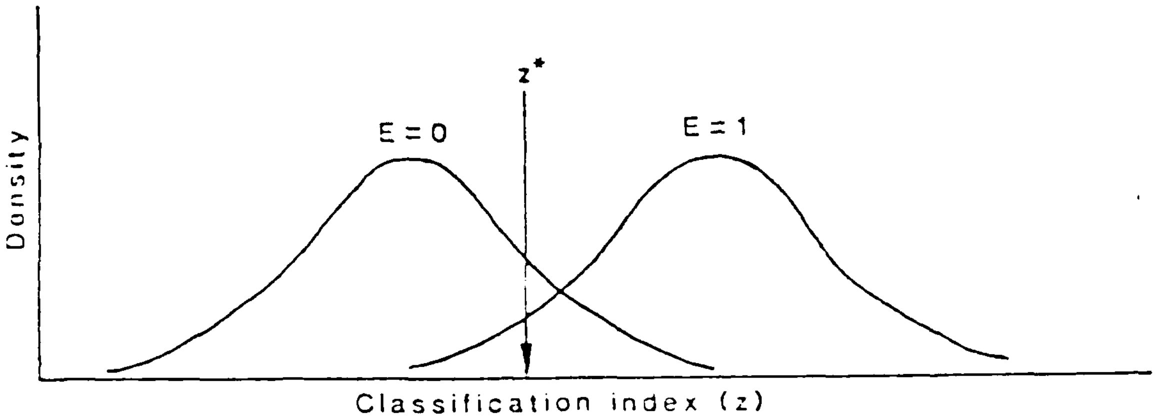
c. Case III: General Solution (Referred to as the Quadratic Model (QUAD) in the text)

$$p(z|E=1) = k/\sigma_1 \exp\{(-1/2)(z-\mu_1)^2/\sigma_1^2\}$$

$$p(z|E=0) = k/\sigma_0 \exp\{(-1/2)(z-\mu_0)^2/\sigma_0^2\}$$

where Λ is the likelihood ratio and $p_0 = p[E=0]$ and $p_1 = p[E=1]$. Thus, the threshold value is

$$z^* = (\mu_0 + \mu_1)/2 + \sigma^2 \ln(p_0/p_1)/(\mu_1 - \mu_0)$$



The position of the threshold depends on the relative values of p_1 and p_0. The threshold moves toward the group with the smallest p_i. If $p_1 = p_0$ the threshold will be the value of z where the densities intersect (i.e., where the densities are equal).

b. Case II: Equal means; different variances

$$\Lambda(z) = \frac{\sigma_0 \exp\{(-1/2)(z-\mu_1)^2/\sigma_1^2\}}{\sigma_1 \exp\{(-1/2)(z-\mu_0)^2/\sigma_0^2\}} \underset{c=0}{\overset{c=1}{\gtrless}} \frac{p_0}{p_1}$$

with the threshold

$$z^* = \pm \left[\frac{2\sigma_0^2\sigma_1^2}{(\sigma_1^2-\sigma_0^2)} \ln\left(\frac{p_0\sigma_1}{p_1\sigma_0}\right) \right]^{1/2}$$

These statements can be combined to give,

$$p(z|E=1)/p(z|E=0) \quad = \quad \Lambda(z) \underset{c=0}{\overset{c=1}{\gtrless}} p[E=0]/p[E=1]$$

Thresholds are the value(s) of z for which

$$\Lambda(z) \quad = \quad p[E=0]/p[E=1]$$

This equation can be solved for z either analytically or numerically depending on the forms of the density functions.

3. Threshold Cases

In order to exemplify the model, the assumption is made that the class conditional distributions are Gaussian. There are essentially three distinct cases that can arise.

a. Case I: Equal variances; different means (Referred to as the equal variance model (EVAR) in the text)

$$p(z|E=1) \quad = \quad k \exp\{(-1/2)(z-\mu_1)^2/\sigma^2\}$$

$$p(z|E=0) \quad = \quad k \exp\{(-1/2)(z-\mu_0)^2/\sigma^2\}$$

where:

$$k \quad = \quad (2\pi)^{-1/2}\sigma^{-1} .$$

$$\Lambda(z) \quad = \quad \frac{\exp\{(-1/2)(z-\mu_1)^2/\sigma^2\}}{\exp\{(-1/2)(z-\mu_0)^2/\sigma^2\}} \underset{c=0}{\overset{c=1}{\gtrless}} \frac{p_0}{p_1}$$

then,

$$p_e = p[E=0] \int_{z \in Z_1} p(z|E=0)dz + p[E=1][1 - \int_{z \in Z_1} p(z|E=1)dz]$$

and algebraic rearrangement yields,

$$p_e = p[E=1] - \int_{z \in Z_1} \{p[E=0]\ p(z|E=0) - p[E=1)\ p(z|E=1)\ dz\}$$

In order to minimize p_e, Z_1 (the decision region for C = 1) will include all those values of z for which the integrand in the expression for p_e will be negative. The decision regions can be symbolically represented as follows:

$$Z_0 = \{z: p[E=0]\ p(z|E=0) - p[E=1]\ p(z|E=1) > 0\}$$

$$Z_1 = \{z: p[E=0]\ p(z|E=0) - p[E=1]\ p(z|E=1) < 0\}$$

An alternative representation is given by,

$$Z_0 = \{z: p[E=0]\ p(z|E=0) > p[E=1]\ p(z|E=1)\}$$

$$= \{z: p[E=0]/p[E=1] > p(z|E=1)/p(z|E=0)\}$$

Likewise,

$$Z_1 = \{z: p[E=0]/p[E=1] < p(z|E=1)/p(z|E=0)\}$$

The decision regions are mutually exclusive and exhaustive (i.e., $Z_0 \cap Z_1 = 0$ and $Z = Z_0 \cup Z_1$).

Thresholds $\equiv$ boundary(s) between decision regions.

$p(z|E=0)$ $\equiv$ class conditional density of z given that E = 0.

$p(z|E=1)$ $\equiv$ class conditional density of z given that E = 1.

$\Lambda(z) = p(z|E=1)/p(z|E=0)$ = the maximum likelihood ratio (i.e., the ratio of class conditional densities).

$p_e = p\{[C=1 \cap E=0] \cup [C=0 \cap E=1]\}$ = the total probability of error.

2. Minimum Probability of Error Criterion

p_e = probability of an incorrect classification.

$$p_e = p[C=1|E=0]\ p[E=0] + p[C=0|E=1]\ p[E=1]$$

where $p[E=1] + p[E=0] = 1$. Note that the events E = 1 and E = 0 are mutually exclusive and exhaustive. The objective is to select decision regions (thresholds) so as to minimize p_e.

$p[C=0|E=1] = \int_{z \in Z_0} p(z|E=1)dz$ = the probability of misclassifying E = 1.

$$p[C=0|E=1] = \int_{z \in Z_0} p(z|E=1)dz + \int_{z \in Z_1} p(z|E=1)dz - \int_{z \in Z_1} p(z|E=1)dz$$

$$p[C=0|E=1] = 1 - \int_{z \in Z_1} p(z|E=1)dz$$

$$p(C=1|E=0] = \int_{z \in Z_1} p(z|E=0)dz$$

these are substituted into the expression for p_e

B. THRESHOLDS (Lowe, 1984a)

1. Notation

$E \equiv$ an event; this is an indicator variable which when $E = 1$, the threatening event occurs, and when $E = 0$, the non-threatening event occurs.

$C \equiv$ the classification of an unknown event which when $C = 1$, the event is classified as a threat, and when $C = 0$, the event is classified as a non-threat.

$P[E = 1] \equiv$ unconditional probability of occurrence of threat.

$P[E = 0] \equiv$ unconditional probability of occurrence of non-threat.

Error of the 1st kind (false alarm) $[C = 1 \cap E = 0]$.

Error of the 2nd kind (miss) $[C = 0 \cap E = 1]$.

$P[C = 1 \cap E = 0] \equiv$ joint probability of an error of the 1st kind.

$P[C = 0 \cap E = 1] \equiv$ joint probability of an error of the 2nd kind.

$P[C = 1 | E = 0] \equiv$ class conditional probability of misclassifying a non-threat.

$P[C = 0 | E = 1] \equiv$ class conditional probability of misclassifying a threat.

$$P[C = 1 \cap E = 0] = P[C = 1 | E = 0]\ P[E = 0].$$

$$P[C = 0 \cap E = 1] = P[C = 0 | E = 1]\ P[E = 0].$$

z = a value of the predictive index (equivalent to y, above).

Z = range of the predictive index on the real line.

For a dichotomous problem, Z is divided into two parts: Z_0, Z_1,

$$C = 0 \text{ if } z \in Z_0$$

$$C = 1 \text{ if } z \in Z_1$$

Independent variable selection for the BMDP9R program begins with a general screening of the entire set of potential predictors. Variables which are identified as redundant, linear combinations of other variables, with respect to the predictand, are deleted from further consideration. The t statistics for the coefficients which minimize the Cp value for each reviewed subset identifies the "best" subset. The number of predictors assigned to each subset can be predefined and for this study each subset equation was required to have six predictors.

The role of regression, once appropriate predictor variables have been selected, is simply that of dimension reduction (representing a multivariate structure by a univariate proxy which constitutes a classificatory or predictive index). This proxy takes the form of a polynomial, linear in its coefficients, of the components of the multivariate structure. The problem now becomes one of determining the form of the state conditional distributions (one for each group of interest; e.g., one, two and three for ceiling categories I, II and III, as used in this study). Once an appropriate form has been selected, it remains, then, to determine the parameters of the class conditional distributions (e.g., means and variances) and then apply an appropriate decision criterion or threshold model.

APPENDIX A

LINEAR REGRESSION AND THRESHOLD MODELS

A. LINEAR REGRESSION

The linear regression techniques used in the study were first presented by Karl (1984) and extended by Diunizio (1984). The least-squares multiple linear regression problem used in the study is the BMDP9R, all possible subsets regression computer program, found in the BMDP Statistical Software Package (University of California, 1983).

The BMDP9R program employs a "best" possible subset, derived independently of variables or variable sequence, calculated from the group of potential predictors. Once this "best" subset is identified, a linear regression equation is fitted to the data, based only upon those selected predictors. The "best" possible subset is identified, a linear regression equation is fitted to the data, based only upon those selected predictors. The "best" possible subset is calculated by a Furnville-Wilson algorithm which provides the user with a variety of subordinate subsets in addition to the "best" subset. Three criteria are available to define the "best" possible subset as a function of independent variables (predictors) and a dependent variable (predictand): the sample R, the adjusted R, and Mallow's Cp. The Mallow's Cp criteria is used in this study, where "best" is defined as the smallest Cp value.

$$dQ/dy(i,j) = [Q(i,j+1) - Q(i,j-1)6/2K$$

where K is the north/south distance between grid points.

5. Develop a scaling process for the predictors, prior to using the cluster analysis, that reduces dimensionality while maintaining the structure of the predictor's characteristics. This may be necessary in view of the widely ranging values of the various MOP's.

6. Use the measures of separability to screen new parameters in order to gain insight into their usefulness without having to make an entire model run.

7. Further pursue the measures of separability combined with cluster analysis as a parameter selection scheme in association with the linear regression models.

8. Develop a system of general "measures of effectiveness" for the MOS project, specifically for those predictands that are categorized and utilize contingency tables. This would provide a means for realistic evaluation of the performance of the various models tested.

an additional parameter. For example, if the particular parameter at TAU-00, TAU-24 and TAU-48 is given the designation S1, S2 and S3, respectively, then the time differencing could be accomplished thus:

TAU-00 forecast period would use forward difference

$$[3 \times S3 - 4 \times S2 + S1]/48$$

TAU-24 forecast period would use centered difference

$$[S3 - 2 \times S2 + S1]/24$$

TAU-48 forecast period would use backward differencing

$$[-S1 + 4 \times S2 - S3]/48$$

These new parameters could then be used as predictors in the models.

4. A new predictor also could be developed at each time period by doing a spatial difference across the observation points to give a representation of advections (i.e., thermal, vorticity, moisture). A potential scheme would be to use a centered difference at the observation position. If the parameter at the observation point was labeled Q(i,j), the east/west advection could be represented by

$$dQ/dX(i,j) = [Q(i+1,j) - Q(i-1,j)]/2L$$

where L is the distance between gridpoints. A north/south advection could be represented by

point out one more significant fact. The low values of separation between category I and II for all the predictors in the area 2 cloud amount study very obviously coincide with the inability of any of the forecast schemes to skillfully forecast category I. This, too, supports the position that new predictors or new combinations of predictors are necessary to improve significantly the results achieved in this study.

It is of interest to note that the most frequently used variables by both the Preisendorfer and regression methods include vorticity (VOR500, VOR925, and DVRTDZ), low level winds (UBLW, U1000), low level vapor pressures (EAIR, E850) and products involving vapor pressure at 700 mb (VE700, TE700).

B. RECOMMENDATIONS

Based on the observations made in this study and the conclusions above, the following recommendations are offered to future researchers:

1. Interpolate the 12 GMT data base to make TAU-00, TAU-24 and TAU-48 MOP's available as predictors at every observation position.

2. Interpolate 00 GMT MOP's to the 1200 GMT ship position to provide 12-hour history as a new predictor.

3. If the parameters described in 1. and 2. above were available, then a time differencing could be done on the predictors to give time trend information to the models as

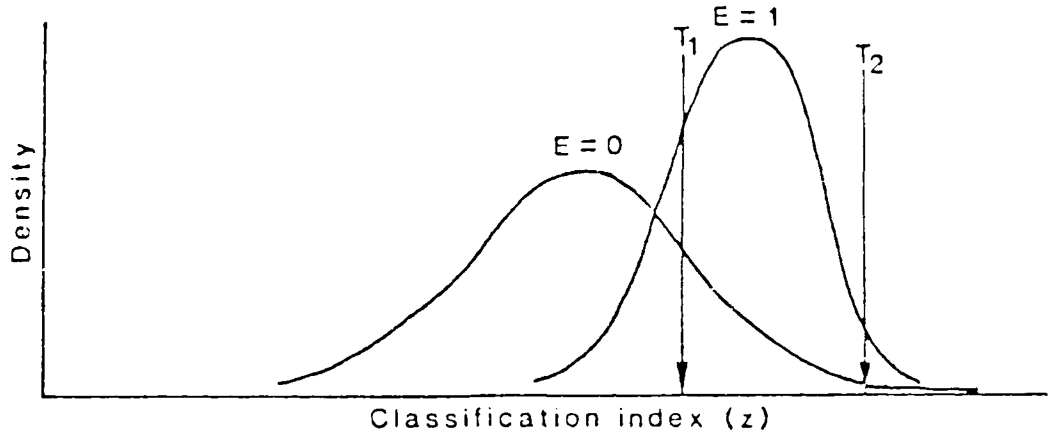
caused both the PR and linear regression methods to gain over 20% in percentage correct, or stated otherwise, they experienced a 40% improvement in their overall percentage correct (i.e., increasing A0 from 47% to 67%). This leads one to believe that the emphasis in further MOS research should not be in pursuing new statistical methods, but rather in pursuing new combinations of old predictors, and new predictors.

The results of the separability measures and cluster analysis individually are not very impressive. The combined use, however, of the techniques with a two-stage regression give the highest A0 for the cloud amount regression schemes, and shows some potential as a predictor selection scheme. The benefits of the two techniques is that it gives the experimenter some control over the parameter selection process, in contrast to the "black box" parameter selection by the BIMED statistical software package. These methods allow the experimenter to adjust the parameter selection according to the category desired to select. For example, if the third category is the most difficult or most desired category for forecasting, then the measures of separability can be used to select predictors providing the maximum separability between the desired categories. These two methods also can provide a screening process for new parameters. With the present models, to evaluate the potential of a single new parameter, the entire model must be run again from the beginning. The results of the measures of separability

proved to be much simpler and less time consuming to apply. The test gives a good first approximation of the significance of any particular model run. The second tool for evaluating the results, the 95% confidence intervals derived from a baseline model contingency table, is very useful in comparing the improvement of each model, and is especially insightful in evaluating degradation of results over the 48-hour time period.

It becomes clear after the first few uses of the various models, that the linear regression techniques are much more easily handled in the developmental stages than the PR or PR+BMD models. When placed into an operational MOS system the PR models will require several orders of magnitude more computer memory storage space than its linear regression counterpart. For these reasons, it would seem that if the PR methods tested here are to be of viable use operationally, they must be able to perform significantly better than the linear regression models.

The results of the ceiling experiment are very encouraging indeed. The first conclusion from this set of experiments is that the premise early in the study that good skill in forecasting cloud amount will be valuable in forecasting ceiling heights is correct. The second conclusion is that the results support the idea that good skill in statistical forecasting of weather elements is more dependent on having good predictors and information than on model type. The addition of a single (perfect) predictor, cloud amount,



The remarks given for the figures in cases I and II are also applicable here. More often than not, only one of a pair of thresholds induced by differing variances will be of real interest. If the variances of the two groups are radically different, then both members of the threshold pair become important.

4. The-Maximum-Likelihood-of-Detection Criteria

For this specific model the following background is provided:

event space: 2 mutually exclusive populations π_0, π_1 forecast decision space: 2 possible forecasts d_0, d_1

d_0 is a correct forecast if π_0 actually occurs
d_1 is a correct forecast if π_1 actually occurs

Problem: select the decision rule d(z) which maps the observation space Z into some forecast space in some optimal manner.

Z may be an observed variable or it may be an univariate index derived from a number of variables.

For this two decision problem, Z is partitioned into two parts, Z_0 and Z_1.

$$d(z) = d_0 \quad \text{if} \quad z \in Z_0$$
$$d(z) = d_1 \quad \text{if} \quad z \in Z_1$$

where $Z_0 \cap Z_1 = 0$ and $Z_0 \cup Z_1 = Z$

The maximum-likelihood-of-detection criteria represents the simplest decision model. The basic involves selecting the forecast (decision) corresponding to the observation (signal) which is the most likely symptom of the event subsequently observed. Consider the following example:

problem: diagnose disease A or disease B.

The observed symptoms occur with probability 0.75 for A and 0.1 for B. By the maximum-likelihood-of-detection criteria (MLDC), diagnose disease A because A is the most likely cause of the observed symptoms (if there is no more information). But if we know that A is rare and B is common, the above decision may not be optimal and MLDC may not be appropriate. MLDC requires only that we know the event conditional probability density functions of the observations. This is:

$p(z|\pi_0)$ and $p(z|\pi_1)$

$$\text{decision rule:} \quad d(z) = \begin{cases} d_1 \text{ if } p(z|\pi_1) > p(z|\pi_0) \\ d_0 \text{ if } p(z|\pi_1) < p(z|\pi_0) \end{cases}$$

In the following development the Gaussian density is used to exemplify the model.

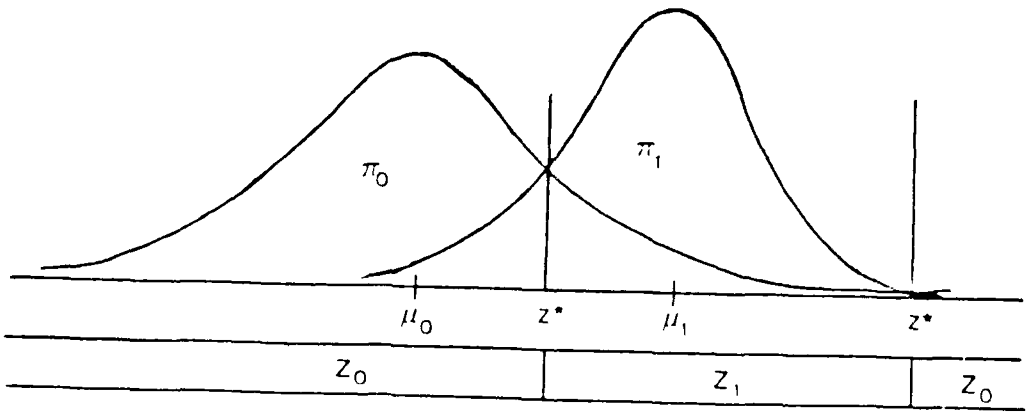
$$p(z|\pi_0) = 1/\sqrt{2\pi\sigma_0} \exp\{-1/2(\frac{z - \bar{z}_0}{\sigma_0})^2\}$$

$$p(z|\pi_1) = 1/\sqrt{2\pi\sigma_1} \exp\{-1/2(\frac{z - \bar{z}_1}{\sigma_1})^2\}$$

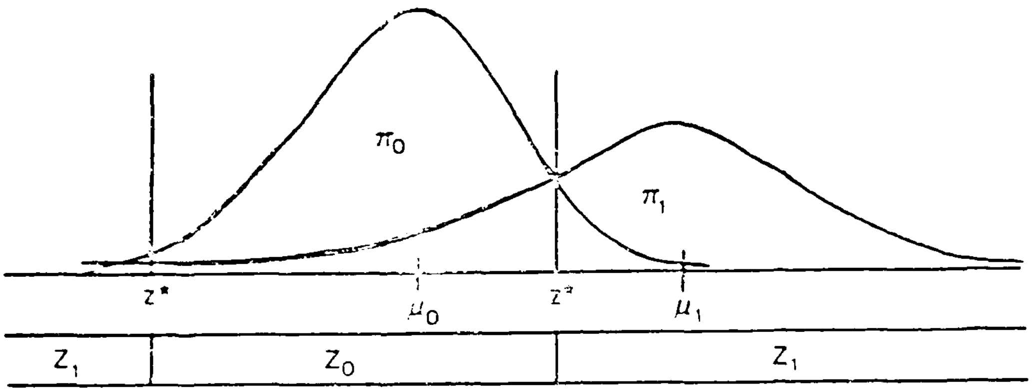
$$\text{definition: likelihood ratio } \Lambda(z) = \frac{p(z|\pi_1)}{p(z|\pi_0)}$$

for convention sake we assume $\bar{z}_1 > \bar{z}_0$,

$\sigma_1^2 > \sigma_0^2$



$\sigma_1^2 > \sigma_0^2$



* note the class having the largest variance has a bifurcated decision region.

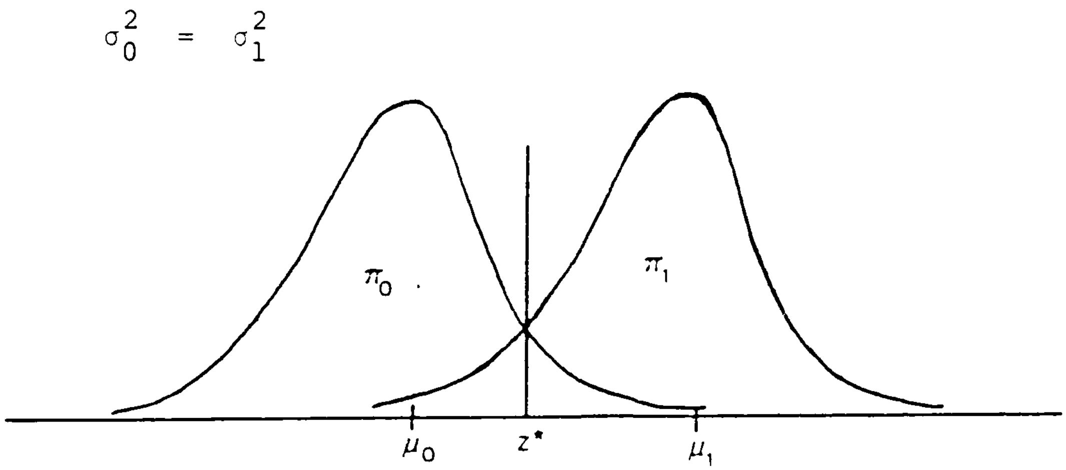
In the case where the variances are equal, the situation simplifies considerably.

$$2\sigma^2(\bar{z}_1 - \bar{z}_0) + \sigma^2(\bar{z}_0^2 - \bar{z}_1^2) \underset{d_0}{\overset{d_1}{\lessgtr}} 0$$

$$2\sigma^2(\bar{z}_1 - \bar{z}_0) - \sigma^2(\bar{z}_1^2 - \bar{z}_0^2) \underset{d_0}{\overset{d_1}{\gtrless}} 0$$

$$2z \underset{d_1}{\overset{d_0}{\gtrless}} (\bar{z}_1 + \bar{z}_0)$$

$$z \gtrless \frac{(\bar{z}_1 + \bar{z}_0)}{2} = z^*$$



It is obvious that z* is simply the average of the means of the class-conditional distributions and is found at the intersections of the two density curves.

In the foregoing, normal class conditional distributions were assumed. This was done because the Gaussian form admits of a rather clean analytical solution. However, the general concept of the minimum probable error decision criteria may be applied to any form of density function. Indeed, the density function of one group need not even be the same form as that for another group (one might be exponential and the other Gaussian). The difficulty with most non-Gaussian forms is that they seldom admit of closed analytical forms and require numerical means in determination of thresholds.

APPENDIX B

MEASURES OF SEPARABILITY

As the testing proceeded through progressive time stages in the study, it became apparent that the methods were unable to separate the categories of scattered and broken clouds, categories I and II. This problem required the investigation of some alternate predictor selection schemes to improve the ability to discriminate between these categories.

The decision information for discriminating between two categories comes from two sources: the separation of the two means and the difference in the variances. The three measures considered in this study are the Divergence, the Bhattacharya distance and the Mahalanobis distance. These three measures attempt to combine both sources of information to come up with a single measure of the ability of a predictor to describe the separation in the categories of the predictand. These measures are applied in the study by stratifying each predictor by event (i.e., predictand category), and calculating the mean and variance of the stratified predictors. Then, for each predictor, the measures of separability are calculated for category I versus II, category I versus III, and category II versus III. The results are shown in tabular form in Tables 11 to 13.

The Mahalanobis distance considers the variances as equal and uses pooled variances of the predictor in the following

univariate form:

$$\frac{(\mu_1 - \mu_2)^2}{\sigma^2}$$

It can be thought of as a signal-to-noise ratio where the difference in the two means is the desired signal, and the noise is the scatter within the whole set (the variance).

The Divergence does not assume equal variance, and, therefore, does not use a pooled variance. It adds to the signal-to-noise ratio two quotients of the variances adjusted by the equal variance value (two). It has the effect of combining the signal-to-noise ratio with information contained in the variances. The Divergence is used in this study in its univariate form:

$$\frac{2(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} + \frac{1}{2}\left(\frac{\sigma_2^2}{\sigma_1^2} + \frac{\sigma_1^2}{\sigma_2^2} - 2\right)$$

The third measure of separability applied to the data set, the Bhattacharyya distance, is a special case of the Chernoff distance. Although more complicated than the Divergence, it also combines the information contained in the mean with that found in the variance. The Bhattacharyya is used in the study in its univariate form as:

$$(\mu_1 - \mu_2)^2 \left(\frac{2}{\sigma_1^2 + \sigma_2^2}\right) + \frac{1}{2} \ln\left(\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1\sigma_2}\right)$$

The results of all three measures of separability applied to the predictors used in the study, are shown in Tables X through XII, for homogeneous area 2 at the time period TAU-00.

APPENDIX C

CLUSTER ANALYSIS

The object of cluster analysis is to take a sample of potential predictor variables of unknown classification and group them into natural classes or clusters. The fact that there is no a priori classification of the sample suggests that cluster analysis is fundamentally a tool for data exploration. That is to say, one wishes to study the data to see if natural and useful groupings do, in fact, exist. It is important to note that for any application of the method there are many possible classifications which can be imposed on a sample. Therefore, the sort of groupings which emerges from an analysis will depend very much on the variables used to represent the predictand. The poor choice of variables can lead to a clustering which is useless for a particular purpose.

The clustering done for cloud amount uses the BMDP Statistical Software (University of California, 1983) P1M program, applied all available Model Output Parameters (MOP's). The P1M provides four measures of similarity (association) for clustering variables and three criteria for linking or combining clusters. Initially, each variable is considered as a separate cluster; then, the two most similar variables are joined to form a cluster. The amalgamating process continues in a stepwise fashion (joining variables or clusters of

variables) until a single cluster is formed that contains all the variables.

As used in this study, the measure of similarity is the absolute value of the correlation. The similarity measure could also be obtained from a measure of the distance, such as the angle between two variables (arccosine of the correlation) or the acute angle corresponding to the arccosine of the absolute value of the correlation.

The linkage rule (the criterion for combining two clusters) can be the minimum distance (or maximum similarity) over all pairings of the variables between the two clusters, the maximum distance (or minimum similarity), or the average distance (or similarity). The average similarity is the arithmetic average of the similarity using all possible pairings of the variables between the two clusters. The maximum similarity (minimum distance), single linkage is used for the MOP's in this study.

The output of the P1M program for homogeneous ocean area 2, at time TAU-00, is:

Predictor Clusters

Cluster	Predictors
1.	D1000, D850, D925, D700, D500, D400, D300, D250
2.	T500, T400, DDDP, T700, T300, TE700
3.	VOR500, VOR925, DVRTDP
4.	TAIR, T1000, T925
5.	EAIR, E1000, E850, E925, EPRD, TE925, E700, E500, T250

6. PBLD, STRTTK, RELH, STRTFQ

7. SMF, SHF

8. BVLW, V850, V925, V1000, V700, V500, V400, V300, V250, VT250, VE700, VT700, UDVDZ

9. UBLW, U850, U925, U1000, U700, U500, U400, U300, U250

10. DRAG, ETRNMT

APPENDIX D

NOGAPS PREDICTOR PARAMETERS AVAILABLE FOR THE NORTH ATLANTIC OCEAN, 15 MAY-15 JULY 1983, EXPERIMENTS

ea: Entire North Atlantic Ocean and Mediterranean Sea

del output time: 1200 GMT (TAU-00)

Model output parameter	Descriptive name of parameter
D1000	1000 mb geopotential height
D925	925 mb geopotential height
D850	850 mb geopotential height
D700	700 mb geopotential height
D500	500 mb geopotential height
D400	400 mb geopotential height
D300	300 mb geopotential height
D250	250 mb geopotential height
TAIR	Surface air temperature
T1000	1000 mb temperature
T925	925 mb temperature
T700	700 mb temperature
T500	500 mb temperature
T400	400 mb temperature
T300	300 mb temperature
T250	250 mb temperature
EAIR	Surface vapor pressure
E1000	1000 mb vapor pressure
E925	925 mb vapor pressure
E850	850 mb vapor pressure
E700	700 mb vapor pressure
E500	500 mb vapor pressure
UBLW	Boundary layer zonal wind component
U1000	1000 mb zonal wind component
U925	925 mb zonal wind component

U850	850 mb zonal wind component
U700	700 mb zonal wind component
U500	500 mb zonal wind component
U400	400 mb zonal wind component
U300	300 mb zonal wind component
U250	250 mb zonal wind component
VBLW	Boundary layer meridional wind component
V1000	1000 mb meridional wind component
V925	925 mb meridional wind component
V850	850 mb meridional wind component
V700	700 mb meridional wind component
V500	500 mb meridional wind component
V400	400 mb meridional wind component
V300	300 mb meridional wind component
V250	250 mb meridional wind component
VOR925	925 mb vorticity
VOR500	500 mb vorticity
PS	Surface pressure
SMF	Surface moisture flux
PBLD	Planetary boundary-layer depth
STRTFQ	Percent stratus frequency
STRTTH	Stratus thickness
SHF	Surface heat flux
ENTRN	Entrainment at top of marine boundary-layer
DRAG	Drag coefficient (C_D)

. Derived parameters

RELH	Surface relative humidity
DVRTDP	Vertical gradient of vorticity (VOR925 - VOR500)
EPRD	Product of vapor pressures (E1000·E850)
DDDP	Height thickness (D925-D250)/675
VT700	Approximation of thermal advection (V700·T700)

UDVDZ	Approximation of thermal advection (U700×(V1000-V500)
TE700	Product of temperature and vapor pressure×(T700 E700)
VT250	Approximation of thermal advection (T250×V250)
TE925	Product of temperature and vapor pressure (T925×e925)

ea: Entire North Atlantic Ocean and Mediterranean Sea

del output time: 1200 GMT (TAU-24 and TAU-48)

Parameters available and derived parameters at TAU-24 and J-48 are the same as those for TAU-00 with the addition of e following five parameters:

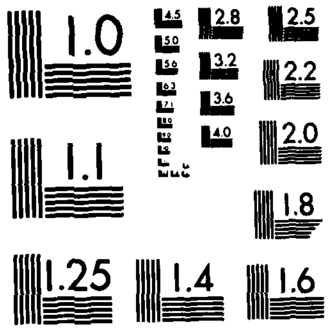
Model output	Descriptive name of parameter
PRECIP	Total amount (mm.) of model precipitation in the last six hours
SHWRS	Total amount (mm.) of model precipitation associated with cumulus convection in the last six hours
INSTAB	Boundary layer inversion instability
DIV925	925 mb Divergence
DIV500	500 mb Divergence

AD-A155 100 AN EVALUATION OF DISCRETIZED CONDITIONAL PROBABILITY AND LINEAR REGRESSIO..(U) NAVAL POSTGRADUATE SCHOOL MONTEREY CA M H WOOSTER DEC 84 2/2

UNCLASSIFIED F/G 4/2 NL

END

FILMED



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

APPENDIX E

VERIFICATION SCORES, DEFINITIONS

FORECAST			
3	R	S	T
2	U	V	W
1	X	Y	Z
	1	2	3
	OBSERVED		

Total = R + S + T + U + V + W + X + Y + Z

A0 = percent correct = (X+V+T)/Total

A1 = one-class error = (U+S+Y+W)/Total

TS1 = Threat score for category I = X/(R+U+X+Y+Z)

TS2 = Threat score for category II = V/(U+V+W+S+Y)

TS3 = Threat score for category III = T/(R+S+T+W+Z)

APPENDIX F

BMDP LINEAR REGRESSION EQUATION PREDICTOR SETS. NORTH ATLANTIC OCEAN (PR+BMD)

These are the derived linear regression equations used as additional predictors in the PR+BMD model. The BMD value of each equation represents an estimate of the category predictand.

I. Area 4, TAU-00, Cloud amount

$$BMD1 = 1.87764 + 0.57546E{-}07 \times U850 + 0.372 \times E700 - 0.4595 \times T500 - 0.00837 \times STRFQ - 9640.3555 \times VOR500 + 28687.457 \times VOR925$$

$$BMD2 = -0.293341 - 0.257147 \times TX + 0.0008191 \times UBLW - 0.055399 \times T500 - 3345.81 \times VOR500 + 8551.9 \times VOR925 - 0.002537 \times EPRD$$

II. Area 2, TAU-00, Cloud amount

$$BMD1 = 2.05292 - 0.09055 \times EAIR + 0.19066E{-}03 \times UBLW - 5335.98438 \times VOR500 + 7474.707 \times VOR925 + 0.00505 \times EPRD + 0.78387E{-}07 \times UDVDZ$$

$$BMD2 = 2.51018 - 0.28119E{-}03 \times U700 + 0.31987 \times E500 + 1.73035 \times DVRTDP + 0.27946E{-}04 \times U700 - 0.2993E{-}05 \times VT250 + 0.00236 \times TE925$$

III. Area 2, TAU-24, Cloud amount

$$BMD1 = 2.98984 - 0.11211 \times EAIR + 0.88063 \times D700$$
$$- 0.01415 \times SHF - 8.28037 \times DIV925$$
$$- 2.27441 \times VOR925 + 0.00656 \times EPRD$$

$$BMD2 = 1.95832 - 0.05608 \times TAIR - 0.084347 \times EAIR$$
$$- 0.01297 \times SMF + 0.12733 \times DIV925$$
$$+ 0.8508E\text{-}05 \times U1000 + 0.192768 \times D700$$

IV. Area 2, TAU-48, Cloud amount

$$BMD1 = 1.5808 + 0.35787E\text{-}03 \times PS - 0.00998 \times EAIR$$
$$- 0.01438 \times V850 + 0.14885E\text{-}03 \times D500$$
$$+ 0.0412 \times V500 + 0.01165 \times TE700$$

$$BMD2 = 2.45617 - 0.5245 \times EAIR + 0.15573 \times E500$$
$$+ 0.06068 \times E925 - 0.2383E\text{-}03 \times STRFQ$$
$$+ 0.0837 \times STRTK - 6.19808 \times DIV925$$

V. Area 2, TAU-00, Ceiling

$$BMD1 = 2.56681 - 0.03478 \times E850 + 0.1884 \times E\text{-}03 \times V700$$
$$- 0.03513 \times T925 + 0.4294E\text{-}03 \times DRAG$$
$$- 2759.3096 \times VOR500 - 0.63253E\text{-}04 \times VE700$$

$$BMD2 = 3.78741 - 0.15939 \times CLAMT - 0.5971E\text{-}04 \times UBLW$$
$$+ 0.1187E\text{-}03 \times V700 - 0.06054 \times E925$$
$$- 2607/6801 \times VOR500 - 0.4057E\text{-}04 \times VE700$$

APPENDIX G

BMDP LINEAR REGRESSION EQUATION PREDICTOR SETS. NORTH ATLANTIC FOR REGRESSION MODELS

These are the derived linear regression equations used in the one and two stage regression models. The BMD value of each equation represents an estimate of the category predictand.

I. Area 4, TAU-00, Cloud amount

a. Two stage regression

$$V1 = 0.763701 - 0.145506 \times TX - 0.004051 \times SHF + 0.13055 \times T1000 + 0.000208 \times UBLW - 0.0001745 \times U1000 + 0.073322 \times E850$$

$$V2 = 0.290741 - 0.12764 \times TX - 0.010425 \times SHF + 0.00017 \times UBLW + 0.11457 \times T1000 - 0.000761 \times U1000 + 0.0792243 \times E850$$

b. Single stage regression

$$V1 = -0.29334 - 0.025715 \times TX + 0.0008191 \times UBLW - 0.055399 \times T500 - 3345.81 \times VOR500 + 8551.8 \times VOR925 - 0.002537 \times EPRD$$

II. Area 2, TAU-00, Cloud amount

a. Single stage regression

$$V1 = 2.05292 - 0.09055 \times EAIR + 0.19066E{-}03 \times UBLW - 5335.9844 \times VOR500 + 7474.707 \times VOR925 + 0.00505 \times EPRD + 0.78387E{-}7 \times UDVDZ$$

b. Single stage regression Separation Test 1

$$
\begin{aligned}
\mathrm{V1} = {} & 1.87611 + 0.98222\mathrm{E}{-04}\times\mathrm{UBLW} - 0.22933\mathrm{E}{-04}\times\mathrm{U1000} \\
& - 0.66081\times\mathrm{E}{-04}\times\mathrm{D850} + .5844\mathrm{E}{-04}\times\mathrm{D700} \\
& - 0.02909\times\mathrm{SHF} + 4128.51953\times\mathrm{VOR925}
\end{aligned}
$$

c. Single stage regression Separation Test 2

$$
\begin{aligned}
\mathrm{V1} = {} & 1.73895 + 0.05728\times\mathrm{E850} + 0.19578\times\mathrm{E700} \\
& - 0.04394\times\llcorner\mathrm{500} + 1.56723\times\mathrm{DVRTDP} \\
& + 0.21519\mathrm{E}{-04}\times\mathrm{VE700} - 0.00557\times\mathrm{TE925}
\end{aligned}
$$

d. Single stage regression Separation & Cluster

$$
\begin{aligned}
\mathrm{V1} = {} & 2.56562 - 0.75713\mathrm{E}{-03}\times\mathrm{PS} - 0.05658\times\mathrm{EAIR} \\
& + 0.49973\mathrm{E}{-04}\times\mathrm{U1000} + 0.07972\times\mathrm{T925} \\
& + 0.00931\times\mathrm{STRTTK} + 4624.29688\times\mathrm{VOR925} \\
& + 0.20189\mathrm{E}{-04}\times\mathrm{VT700} + 0.00127\times\mathrm{TE700}
\end{aligned}
$$

e. Two stage regression separation test

$$
\begin{aligned}
\mathrm{V1} = {} & 1.3457 + 0.43966\mathrm{E}{-04}\times\mathrm{UBLW} + 0.5055\mathrm{E}{-05}\times\mathrm{D1000} \\
& + 0.64366\mathrm{E}{-04}\times\mathrm{U1000} - 0.61831\mathrm{E}{-05}\times\mathrm{D850} \\
& + 0.00285\times\mathrm{SHF} + 2488.08984\times\mathrm{VOR925}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{V2} = {} & 2.3040 - 0.01869\times\mathrm{E850} + 0.13171\times\mathrm{E700} \\
& - 0.06215\times\mathrm{E500} + 1.46893\times\mathrm{DVRTDP} \\
& + 0.39909\mathrm{e}{-05}\times\mathrm{VE700} - 0.2294\mathrm{e}{-04}\times\mathrm{TE925}
\end{aligned}
$$

III. Area 2, TAU-24, Cloud Amount

a. Single stage regression

$$
\begin{aligned}
\mathrm{V1} = {} & 2.98984 - 0.11211\times\mathrm{EAIR} + 0.88063\times\mathrm{D700} \\
& - 0.01415\times\mathrm{SHF} - 8.28037\times\mathrm{DIV925} \\
& - 2.27441\times\mathrm{VOR925} + 0.00656\times\mathrm{EPRD}
\end{aligned}
$$

IV. Area 2, TAU-48, Cloud Amount

a. Single stage regression

$$V1 = 2.45617 - 0.05245 \times EAIR + 0.15573 \times E500$$
$$+ 0.06068 \times E925 - 0.2383E-03 \times STRFQ$$
$$+ 0.00837 \times STRTK - 6.19809 \times DIV925$$

V. Area 2, TAU-00, Ceiling

a. Single stage regression--no cloud amount variable

$$V1 = 2.56681 - 0.03478 \times E950 + 0.18836E-03 \times V700$$
$$- 0.03513 \times T925 + 0.4294E-03 \times DRAG$$
$$- 2759.3096 \times VOR500 - 0.63253E-04 \times VE700$$

b. Single stage regression with cloud amount variable

$$V1 = 3.78741 - 0.15939 \times CLAMT - 0.5971E-04 \times UBLW$$
$$+ 0.1187E-03 \times V700 - 0.06054 \times E925$$
$$- 2607/6801 \times VOR500 - 0.4057E-04 \times VE700$$

APPENDIX H

TABLES

TABLE I

A summary of 1200 GMT cloud amount observations, 15 May to 07 July 1983, North Atlantic Ocean homogeneous areas as shown in Fig. 1: TAU-00

Area	Total	CAT I	CAT II	CAT III
All	11428	4022 (.35)	4485 (.39)	2921 (.26)
1	1686	297 (.18)	675 (.40)	714 (.42)
2	1766	359 (.20)	704 (.40)	672 (.38)
3	1129	324 (.29)	355 (.31)	450 (.40)
4	3182	1019 (.32)	1286 (.40)	877 (.28)
5	2425	1259 (.52)	868 (.36)	298 (.12)
6	2067	842 (.41)	829 (.40)	396 (.19)
7	564	157 (.28)	280 (.50)	127 (.09)
8	1015	522 (.52)	402 (.40)	91 (.09)

TABLE II

A summary of 1200 GMT cloud amount observations, 15 May to 07 July 1983, North Atlantic Ocean homogeneous areas as shown in Fig. 1: TAU-24

Area	Total	CAT I	CAT II	CAT III
All	9416	3378 (.36)	3616 (.38)	2422 (.26)
1	1460	281 (.19)	583 (.40)	596 (.41)
2	1422	290 (.20)	550 (.39)	582 (.41)
3	916	259 (.28)	298 (.33)	359 (.39)
4	2592	857 (.33)	988 (.38)	747 (.29)
5	1992	1050 (.53)	719 (.36)	223 (.11)
6	1684	690 (.41)	652 (.39)	342 (.20)
7	458	140 (.31)	207 (.45)	111 (.24)
8	874	457 (.52)	351 (.40)	66 (.08)

TABLE III

A summary of 1200 GMT cloud amount observations, 15 May to 07 July 1983, North Atlantic Ocean homogeneous areas as shown in Fig. 1: TAU-48

Area	Total	CAT I	CAT II	CAT III
All	10775	3817 (.35)	4150 (.39)	2808 (.26)
1	1676	339 (.20)	691 (.41)	646 (.39)
2	1644	327 (.20)	656 (.40)	661 (.40)
3	1046	308 (.29)	336 (.32)	402 (.38)
4	2976	947 (.32)	1145 (.38)	884 (.30)
5	2264	1166 (.52)	823 (.36)	275 (.12)
6	1949	804 (.41)	753 (.39)	392 (.20)
7	524	158 (.30)	234 (.45)	132 (.25)
8	976	505 (.52)	382 (.39)	89 (.09)

TABLE IV

A summary of 1200 GMT ceiling observations, 15 May to 07 July 1983, North Atlantic Ocean homogeneous area 2 for TAU-00

Area	Total	CAT I	CAT II	CAT III
2	1791	415 (.23)	672 (.38)	704 (.39)

TABLE V

Summary of the cloud amount contingency table statistical results for all models used in the North Atlantic homogeneous area 4 evaluated at the FATJUNE 1983 TAU-00 time period

Model		Dependent A0(%)	Dependent TS1	Dependent TS2	Dependent TS3	Independent A0(%)	Independent TS1	Independent TS2	Independent TS3
BMD TS	EVAR	44.47	.18	.39	.11	45.27	.18	.40	.11
	QUAD	44.47	.18	.39	.11	45.27	.18	.40	.11
	MLDC	45.78	.36	.30	.19	43.30	.33	.27	.19
BMD SS	EVAR	47.70	.31	.34	.26	46.20	.28	.33	.27
	QUAD	47.98	.30	.36	.24	46.67	.27	.35	.25
	MLDC	47.00	.36	.23	.34	43.67	.31	.20	.34
PR	MP1	90.02	.81	.82	.82	44.14	.24	.30	.30
	MP2	90.02	.83	.82	.80	41.14	.26	.27	.25
	NATR	87.39	.79	.74	.81	45.36	.22	.34	.29
PR+BMD	MP1	88.43	.80	.78	.80	45.27	.27	.34	.23
	MP2	88.43	.82	.78	.78	45.92	.29	.35	.22
	NATR	86.50	.79	.72	.79	44.14	.17	.39	.12

TABLE VI

Summary of the cloud amount contingency table statistical results for all models used in the North Atlantic homogeneous area 2 evaluated at the FATJUNE 1983 TAU-00 time period

Model		Dependent				Independent			
		A0(%)	TS1	TS2	TS3	A0(%)	TS1	TS2	TS3
BMD SS	EVAR	54.02	.06	.39	.45	49.41	.09	.36	.39
	QUAD	53.76	.05	.41	.42	49.02	.08	.38	.37
	MLDC	51.76	.27	.30	.45	47.40	.26	.26	.39
PR	MP1	85.43	.63	.73	.81	51.26	.11	.38	.40
	MP2	85.43	.68	.74	.79	51.26	.11	.38	.40
	NATR	83.33	.61	.68	.80	48.07	.08	.42	.30
PR+BMD	MP1	91.12	.78	.83	.87	51.26	.17	.35	.43
	MP2	91.12	.80	.83	.86	50.08	.13	.36	.41
	NATR	89.78	.78	.78	.86	53.43	.11	.42	.40
SEPER	1vs2	50.84	.00	.38	.41	45.39	.00	.34	.34
SEPER	2vs3	53.60	.00	.41	.39	48.91	.00	.36	.40
SEPER + CLUSTER		49.50	.00	.37	.39	44.05	.00	.34	.31
SEPER TS	EVAR	53.85	.05	.41	.43	50.59	.05	.39	.39
	QUAD	53.18	.03	.42	.41	50.08	.01	.39	.39

DEPENDENT DATA

FORECAST	OBSERVED 1	OBSERVED 2	OBSERVED 3
3	17	47	75
2	521	735	502
1	139	71	27

A0(%): 44.47

TS1 : .18

TS2 : .39

TS3 : .11

INDEPENDENT DATA

FORECAST	OBSERVED 1	OBSERVED 2	OBSERVED 3
3	8	22	37
2	263	375	240
1	71	36	15

A0(%): 45.27

TS1 : .18

TS2 : .40

TS3 : .11

Figure 5. Contingency table results for the area 4, TAU-00, two-stage regression, QUAD model for cloud amount

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	17	47	75
2	521	735	502
1	139	71	27

AO(%): 44.47

TS1 : .18

TS2 : .39

TS3 : .11

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	8	22	37
2	263	375	240
1	71	36	15

AO(%): 45.27

TS1 : .18

TS2 : .40

TS3 : .11

Figure 4. Contingency table results for the area 4, TAU-00, two-stage regression, EVAR model for cloud amount

SIGNIFICANCE TEST

Null Hypothesis Contingency

Table (Chance)

FORECAST \ OBSERVED	1	2	3
3	114	145	98
2	114	144	97
1	114	144	97

A0(%): 33.36

TS1 .20

TS2 : .22

TS3 : .18

Area 4 TAU00

95% CONFIDENCE INTERVAL

A0(%). 30.53 - 36.19

TS1 .17 - .22

TS2 : .20 - .25

TS3 .15 - .20

Figure 3. Confidence intervals for significance with respect to chance--area 4, TAU-00, cloud amount

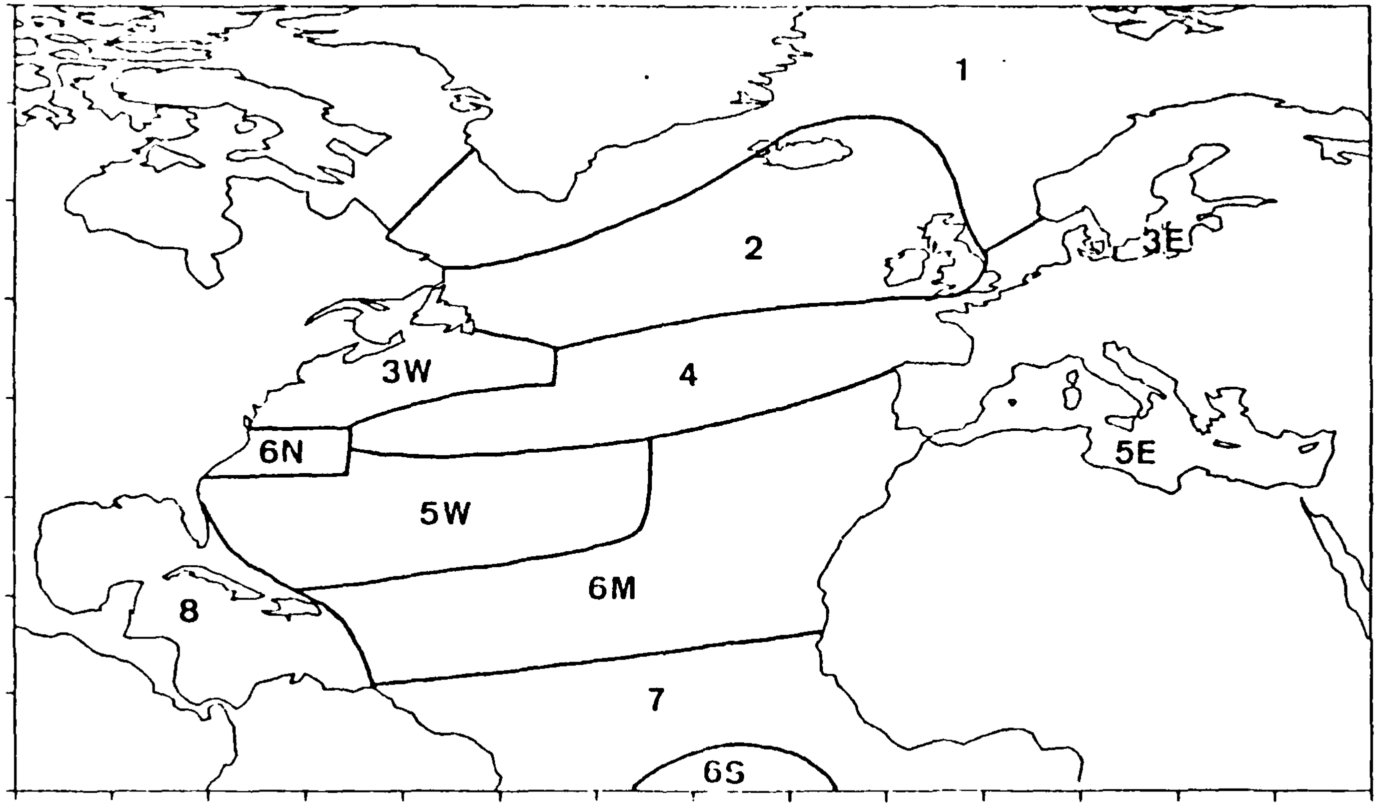


Figure 2. Homogeneous areas for the North Atlantic ocean, June and July, from Lowe (1984b)

APPENDIX I

FIGURES

OPERATIONAL U.S. NAVY
MODEL OUTPUT STATISTICS (MOS)
DEVELOPMENT SCHEDULE

<table>
<tr><th>ELEMENT GROUP / OCEAN BASIN</th><th>VISIBILITY
CEILING
CLOUD AMOUNT
OBSTR TO VIS</th><th>SIG WAVE HT
WIND WAVES
OCEAN SWELL</th><th>SFC WIND
SFC TEMP
SFC DEWPOINT</th><th>PRECIP OCCUR
PRECIP AMT
TRW OCCUR</th></tr>
<tr><td>NORTH ATLANTIC</td><td>1987</td><td>1988</td><td rowspan="3">1990</td><td rowspan="3">1991</td></tr>
<tr><td>MEDITERRANEAN</td><td rowspan="2">1988</td><td rowspan="2">1989</td></tr>
<tr><td>NORTH PACIFIC</td></tr>
<tr><td>INDIAN</td><td rowspan="3">1989</td><td rowspan="3">1990</td><td rowspan="3">1991</td><td rowspan="3">1992</td></tr>
<tr><td>SOUTH ATLANTIC</td></tr>
<tr><td>SOUTH PACIFIC</td></tr>
</table>

Figure 1. Proposed U.S. Navy model output statistics (MOS) development schedule

TABLE XII (Continued)

VARIABLE	BHATTACHARYYA	DIVERGENCE	MAHALANOBIS
T300	0.21056	0.22100	0.20881
U300	0.01408	0.02348	0.01272
V300	0.07071	0.07155	0.07062
D925	0.00771	0.00836	0.00762
T925	0.20461	0.21458	0.20294
E925	0.20111	0.30206	0.18641
U925	0.00317	0.01440	0.00158
V925	0.12695	0.13894	0.12507
D250	0.19249	0.19258	0.19250
T250	0.01740	0.02707	0.01600
U250	0.01320	0.02348	0.01172
V250	0.06705	0.06854	0.06687
PBLD	0.01446	0.01503	0.01438
STFQ	0.06272	0.06298	0.06267
STSK	0.01442	0.01972	0.01367
SHF	0.36083	0.36084	0.36084
ETRN	0.01873	0.09273	0.00844
DRAG	0.08309	0.22548	0.06334
VOR5	0.13266	0.13638	0.13223
VOR9	0.00003	0.00005	0.00003
RHSU	0.01062	0.02140	0.00910
DDDP	0.27874	0.28099	0.27825
DVRT	0.35180	0.36540	0.34936
EPRD	0.21446	0.34633	0.19553
VT70	0.02794	0.21941	0.00217
VE70	0.17673	0.46591	0.13800
UDVZ	0.00548	0.03222	0.00169
TE70	0.18493	0.32513	0.16501
VT25	0.06871	0.07051	0.06849
TE92	0.24765	0.39902	0.22596
RH50	0.12367	0.12393	0.12361

TABLE XII

Listing of the measures of separability, by predictor, for cloud amount categories II versus III, North Atlantic Ocean homogeneous area 2 at time period TAU-00

VARIABLE	BHATTACHARYYA	DIVERGENCE	MAHALANOBIS
PS	0.00117	0.00218	0.00103
TX	0.12979	0.13784	0.12850
EX	0.07575	0.09323	0.07315
SMF	0.20953	0.22308	0.20791
UBLW	0.00386	0.02460	0.00092
VBLW	0.12082	0.14304	0.11745
D100	0.00137	0.00229	0.00124
T100	0.12849	0.13413	0.12757
E100	0.10728	0.14164	0.10219
U100	0.00279	0.01875	0.00052
V100	0.10973	0.13279	0.10625
D850	0.02216	0.02261	0.02210
E850	0.30505	0.44064	0.028520
U850	0.00403	0.01390	0.00262
V850	0.12095	0.12695	0.11997
D700	0.07096	0.07210	0.07083
T700	0.31687	0.32132	0.31640
E700	0.50219	0.59661	0.48725
U700	0.00688	0.01573	0.00561
V700	0.08627	0.08864	0.08588
D500	0.14117	0.14293	0.14100
T500	0.22404	0.22409	0.24405
E500	0.34663	0.41264	0.33632
U500	0.01009	0.01637	0.00919
V500	0.07677	0.07727	0.07667
D400	0.16929	0.17051	0.16919
T400	0.22131	0.22215	0.22110
U400	0.01272	0.01996	0.01168
V400	0.07573	0.07575	0.07573
D300	0.18786	0.18828	0.18785

TABLE XI (Continued)

VARIABLE	BHATTACHARYYA	DIVERGENCE	MAHALANOBIS
T300	0.20562	0.24049	0.19159
U300	0.03037	0.03054	0.03045
V300	0.16807	0.16866	0.16694
D925	0.04597	0.04739	0.04533
T925	0.06919	0.07071	0.06829
E925	0.14805	0.25984	0.12245
U925	0.06896	0.09353	0.06296
V925	0.21093	0.26951	0.19104
D250	0.04663	0.04671	0.04651
T250	0.10538	0.12246	0.09963
U250	0.02878	0.02899	0.02886
V250	0.16161	0.16237	0.16037
PBLD	0.09085	0.09125	0.09032
STFQ	0.15207	0.15208	0.15220
STSK	0.09857	0.09902	0.09797
SHF	0.28824	0.35556	0.26172
ETRN	0.01747	0.04776	0.01380
DRAG	0.06575	0.26496	0.03517
VOR5	0.01224	0.01355	0.01195
VOR9	0.12924	0.14618	0.12275
RHSU	0.06157	0.06976	0.06183
DDDP	0.15300	0.15966	0.14893
DVRT	0.27339	0.27967	0.26707
EPRD	0.13559	0.24367	0.11154
VT70	0.00997	0.07588	0.00080
VE70	0.28308	0.76258	0.19313
UDVZ	0.01686	0.05881	0.01041
TE70	0.04632	0.15561	0.02886
VT25	0.17008	0.17019	0.16961
TE92	0.12640	0.21771	0.10574
RH50	0.28494	0.30526	0.27214

TABLE XI

Listing of the measures of separability, by predictor, for cloud amount categories I versus III, North Atlantic Ocean homogeneous area 2 at time period TAU-00

VARIABLE	BHATTACHARYYA	DIVERGENCE	MAHALANOBIS
PS	0.08177	0.14363	0.07800
TX	0.04016	0.04056	0.03990
EX	0.00637	0.01161	0.00552
SMF	0.13855	0.13888	0.13915
UBLW	0.08358	0.15122	0.06957
VBLW	0.21032	0.30548	0.18303
D100	0.06498	0.06676	0.06403
T100	0.07172	0.07229	0.07120
E100	0.04262	0.07060	0.03708
U100	0.07400	0.14819	0.05957
V100	0.18603	0.28651	0.15951
D850	0.02559	0.02626	0.02532
E850	0.29357	0.45542	0.24720
U850	0.06046	0.07166	0.05731
V850	0.20114	0.23462	0.18768
D700	0.00182	0.00210	0.00178
T700	0.11150	0.12273	0.11297
E700	0.65337	0.77698	0.58537
U700	0.04171	0.04595	0.04043
V700	0.15351	0.16135	0.14901
D500	0.00839	0.01025	0.00821
T500	0.09364	0.09505	0.09255
E500	0.36905	0.51689	0.31872
U500	0.03106	0.03106	0.03106
V500	0.15269	0.15338	0.15157
D400	0.02288	0.02363	0.02294
T400	0.13030	0.14202	0.12517
U400	0.03115	0.03k38	0.03124
V400	0.16649	0.16697	0.16550
D300	0.03975	0.03976	0.03978

TABLE X (Continued)

VARIABLE	BHATTACHARYYA	DIVERGENCE	MAHALANOBIS
T300	0.00155	0.00861	0.00053
U300	0.00636	0.01845	0.00477
V300	0.01923	0.02207	0.01858
D925	0.09076	0.09473	0.08882
T925	0.03484	0.03853	0.03483
E925	0.00652	0.00678	0.00646
U925	0.05098	0.05353	0.04999
V925	0.00962	0.02688	0.00695
D250	0.05071	0.05106	0.05043
T250	0.04040	0.04144	0.03994
U250	0.00656	0.02001	0.00478
V250	0.01878	0.02316	0.01786
PBLD	0.03187	0.03379	0.03126
STFQ	0.02029	0.02066	0.02033
STSK	0.03631	0.04517	0.03426
SHF	0.02615	0.09518	0.01552
ETRN	0.00162	0.01092	0.00029
DRAG	0,00550	0.00977	0.00482
VOR5	0.06948	0.07893	0.06654
VOR9	0.12475	0.14278	0.11831
RHSU	0.02086	0.02103	0.02076
DDDP	0.02176	0.02293	0.02142
DVRT	0.00438	0.00577	0.00422
EPRD	0.01150	0.01259	0.01144
VT70	0.00871	0.03907	0.00423
VE70	0.01574	0.03484	0.01260
UDVZ	0.00476	0.00643	0.00448
TE70	0.06954	0.07131	0.07002
VT25	0.02086	0.02366	0.02020
TE92	0.02343	0.03042	0.02291
RG50	0.02931	0.04528	0.02623

TABLE X

Listing of the measures of separability, by predictor, for cloud amount categories I versus II, North Atlantic Ocean homogeneous area 2 at time period TAU-00

VARIABLE	BHATTACHARYYA	DIVERGENCE	MAHALANOBIS
PS	0.09467	0.14153	0.09295
TX	0.02457	0.02943	0.02429
EX	0.04045	0.04401	0.04054
SMF	0.01079	0.02040	0.00920
UBLW	0.06747	0.08060	0.06381
VBLW	0.01148	0.03612	0.00770
D100	0.08318	0.08845	0.09098
T100	0.00768	0.01029	0.00740
E100	0.01717	0.01749	0.01720
U100	0.06117	0.08200	0.05623
V100	0.00962	0.03601	0.00566
D850	0.09544	0.09767	0.09403
E850	0.00064	0.00169	0.00048
U850	0.03957	0.03961	0.03950
V850	0.00910	0.02012	0.00735
D700	0.09371	0.09400	0.09329
T700	0.04722	0.04876	0.04746
E700	0.01074	0.01257	0.01037
U700	0.01787	0.01871	0.01788
V700	0.00922	0.01080	0.00891
D500	0.07878	0.07878	0.07881
T500	0.03095	0.03293	0.03033
E500	0.00232	0.01756	0.00015
U500	0.00822	0.01458	0.00746
V500	0.01307	0.01308	0.01305
D400	0.06663	0.06669	0.06650
T400	0.01693	0.02318	0.01573
U400	0.00697	0.01702	0.00568
V400	0.01696	0.01763	0.01676
D300	0.05530	0.05561	0.05502

TABLE IX

Summary of the ceiling contingency table statistical results for all models used in the North Atlantic homogeneous area 2 evaluated at the FATJUNE 1983 TAU-00 time period

Model		Dependent				Independent			
		AO(%)	TS1	TS2	TS3	AO(%)	TS1	TS2	TS3
BMD SS	EVAR	47.24	.26	.43	.00	44.56	.22	.42	.00
	QUAD	-	-	-	-	-	-	-	-
	MLDC	46.98	.26	.40	.13	45.56	.22	.39	.15
PR+BMD1	MP1	59.13	.35	.36	.49	47.24	.28	.28	.36
	MP2	59.13	.39	.38	.47	45.73	.30	.24	.34
	NATR	46.57	.16	.35	.36	46.06	.18	.38	.23
PR+CLT	MP1	91.05	.73	.82	.91	68.60	.24	.49	.72
	MP2	91.05	.77	.81	.91	68.60	.24	.49	.72
	NATR	89.60	.72	.77	.91	63.36	.20	.45	.64
BMD SS + CLT	EVAR	65.91	.35	.51	.60	67.50	.35	.52	.62
	QUAD	65.49	.33	.51	.59	67.67	.34	.53	.62

TABLE VIII

Summary of the cloud amount contingency table statistical results for all models used in the North Atlantic homogeneous area 2 evaluated at the FATJUNE 1983 TAU-48 time period

Model		Dependent				Independent			
		A0(%)	TS1	TS2	TS3	A0(%)	TS1	TS2	TS3
BMD SS	EVAR	48.68	.06	.35	.39	45.32	.04	.32	.36
	QUAD	-	-	-	-	-	-	-	-
	MLDC	45.96	.16	.28	.39	42.93	.15	.25	.36
PR	MP1	93.03	.85	.87	.88	46.47	.14	.26	.40
	MP2	93.03	.85	.87	.88	45.76	.09	.32	.35
	NATR	91.71	.84	.82	.88	43.99	.10	.29	.36
PR+BMD	MP1	93.03	.84	.86	.89	49.47	.00	.34	.42
	MP2	93.03	.85	.87	.88	49.47	.00	.34	.42
	NATR	91.53	.85	.81	.88	45.23	.20	.29	.34

TABLE VII

Summary of the cloud amount contingency table statistical results for all models used in the North Atlantic homogeneous area 2 evaluated at the FATJUNE 1983 TAU-24 time period

Model		Dependent				Independent			
		A0(%)	TS1	TS2	TS3	A0(%)	TS1	TS2	TS3
BMD SS	EVAR	47.92	.02	.32	.41	46.97	.01	.33	.40
	QUAD	48.11	.02	.36	.38	46.59	.01	.35	.36
	MLDC	43.94	.20	.20	.41	42.05	.17	.19	.40
PR	MP1	65.48	.30	.44	.59	50.31	.00	.34	.44
	MP2	65.48	.38	.47	.56	49.49	.00	.30	.45
	NATR	60.29	.21	.42	.53	45.62	.05	.32	.39
PR+BMD	MP1	91.85	.81	.84	.87	49.69	.13	.33	.42
	MP2	91.85	.81	.84	.87	49.69	.17	.31	.42
	NATR	90.84	.81	.79	.88	46.64	.14	.31	.37

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	48	102	144
2	192	396	282
1	437	355	178

A0(%): 45.78

TS1 : .36

TS2 : .30

TS3 : .19

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	28	51	72
2	109	185	134
1	205	197	86

A0(%): 43.30

TS1 : .33

TS2 : .27

TS3 : .19

Figure 6. Contingency table results for the area 4, TAU-00, two-stage regression, MLDC model for cloud amount

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	41	122	198
2	342	526	333
1	292	205	73

A0(%): 47.70

TS1 : .31

TS2 : .34

TS3 : .26

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	19	53	97
2	186	259	163
1	137	121	32

A0(%): 46.20

TS1 : .28

TS2 : .33

TS3 : .27

Figure 7. Contingency table results for the area 4, TAU-00, single-stage regression, EVAR model for cloud amount

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	31	104	181
2	366	563	353
1	280	186	70

A0(%): 47.98

TS1 : .30

TS2 : .36

TS3 : .24

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	16	43	87
2	195	280	175
1	131	110	30

A0(%): 46.67

TS1 : .27

TS2 : .35

TS3 : .25

Figure 8. Contingency table results for the area 4, TAU-00, single-stage regression, QUAD model for cloud amount

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	126	298	348
2	156	260	144
1	395	295	112

A0(%): 47.00

TS1 : .36

TS2 : .23

TS3 : .34

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	60	144	170
2	107	121	67
1	175	168	55

A0(%): 43.67

TS1 : .31

TS2 : .20

TS3 : .34

Figure 9. Contingency table results for the area 4, TAU-00, single-stage regression, MLDC model for cloud amount

BASELINE INTERVAL

Baseline model : Area 4 TAU00 BMD-EVAR

FORECAST \ OBSERVED	1	2	3
3	19	53	97
2	186	259	163
1	137	121	32

A0(%) 46.20

TS1 : .28

TS2 .33

TS3 .27

BASELINE CONFIDENCE INTERVAL

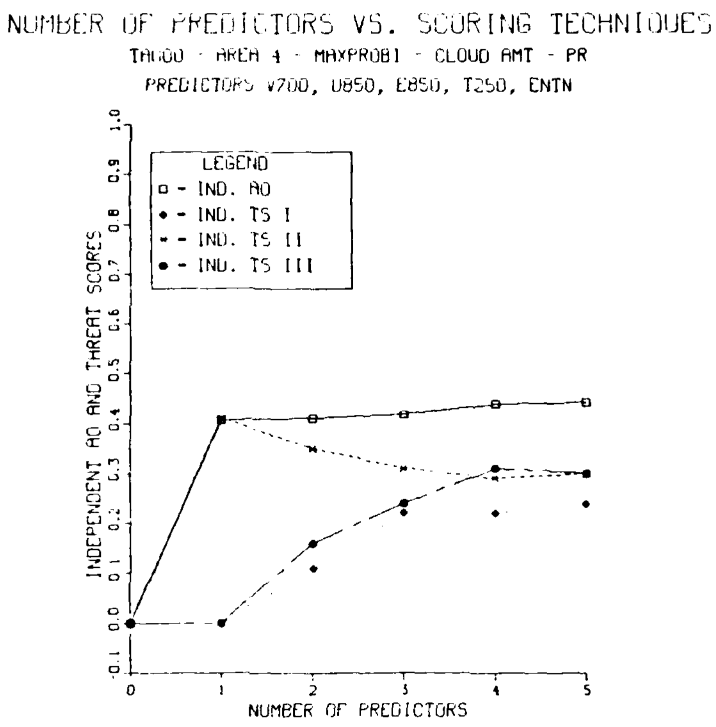
A0(%) 43.21 - 49.19

TS1 .25 - .31

TS2 .30 - .36

TS3 .24 - .30

Figure 10. Confidence intervals for significance with respect to baseline--area 4, TAU-00, cloud amount



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	41	81	659
2	64	763	12
1	499	9	6

AO(%): 90.02

TS1 : .83

TS2 : .82

TS3 : .80

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	68	132	164
2	119	202	123
1	105	99	55

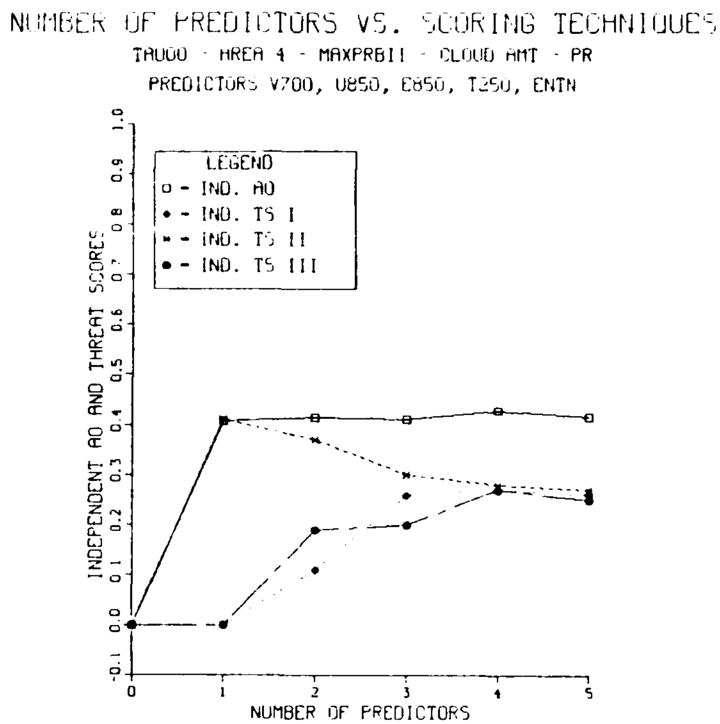
AO(%): 44.14

TS1 : .24

TS2 : .30

TS3 : .30

Figure 11a. Cloud amount skill diagram and contingency table results for area 4, TAU-00 (PR model) for the MAXPROB I strategies. The contingency table corresponds to those scores associated with the maximum independent AO achieved at the fifth predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	5	10	554
2	12	780	81
1	587	63	42

A0(%): 90.02

TS1 : .83

TS2 : .82

TS3 : .80

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	43	89	118
2	111	183	143
1	138	161	79

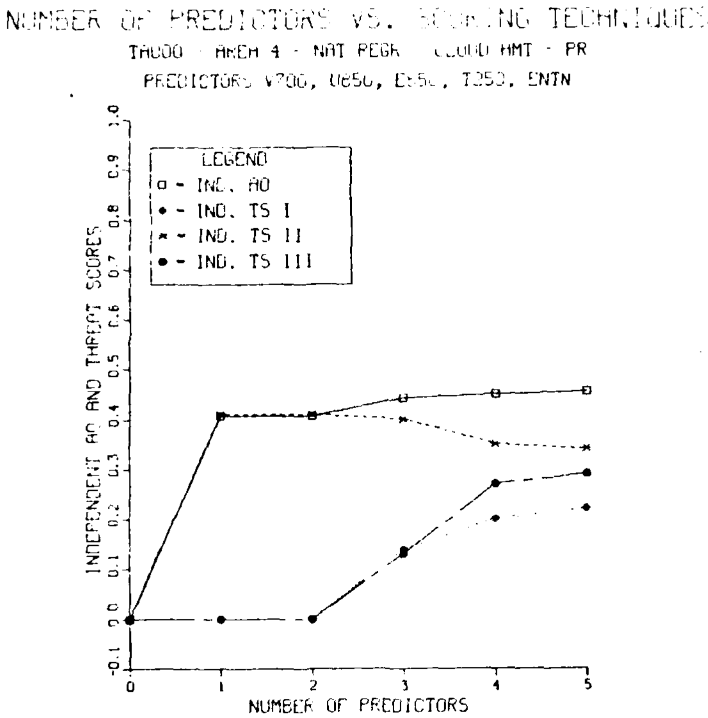
A0(%): 41.14

TS1 : .26

TS2 : .27

TS3 : .25

Figure 11b. Cloud amount skill diagram and contingency table results for area 4, TAU-00 (PR model) for the MAXPROB II strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the fifth predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	0	77	612
2	118	767	65
1	486	9	0

A0(%): 87.39

TS1 : .79

TS2 : .74

TS3 : .81

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	48	97	139
2	151	252	156
1	93	84	47

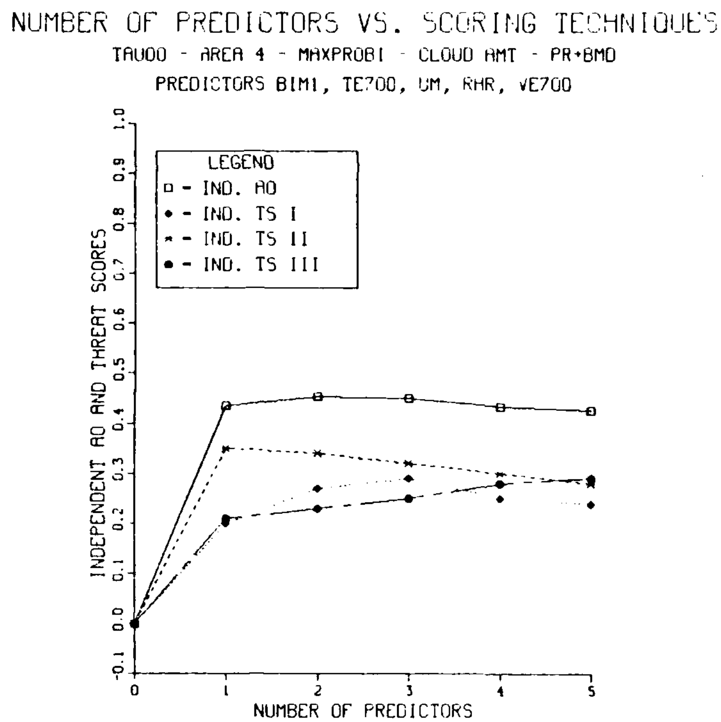
A0(%): 45.36

TS1 : .22

TS2 : .34

TS3 : .29

Figure 11c. Cloud amount skill diagram and contingency table results for area 4, *TAU-00* (PR model) for the natural regression strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the fifth predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	29	98	587
2	87	739	11
1	561	16	6

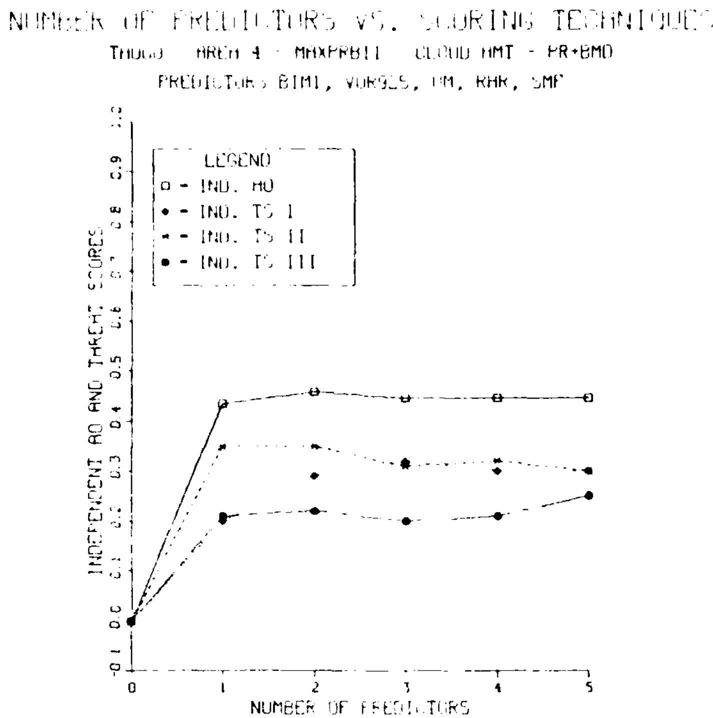
A0(%): 88.43
TS1 : .80
TS2 : .78
TS3 : .80

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	43	73	94
2	168	258	157
1	131	102	41

A0(%): 45.27
TS1 : .27
TS2 : .34
TS3 : .23

Figure 12a. Cloud amount skill diagram and contingency table results for area 4, TAU-00 (PR+BMD model) for the MAXPROB I strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the second predictor.



DEPENDENT DATA

FORECAST	1	2	3
3	9	15	491
2	15	783	87
1	653	95	6

OBSERVED

A0(%) 88.43

TS1 : .82

TS2 : .78

TS3 : .78

INDEPENDENT DATA

FORECAST	1	2	3
3	43	69	88
2	157	260	157
1	142	104	47

OBSERVED

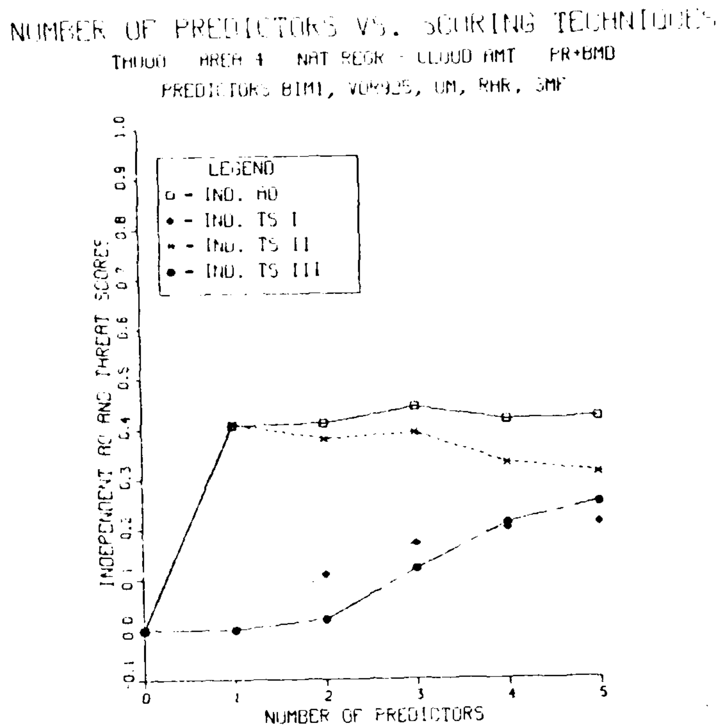
A0(%) 45.92

TS1 : .29

TS2 : .35

TS3 : .22

Figure 12b. Cloud amount skill diagram and contingency table results for area 4, TAU-00 (PR+BMD model) for the MAXPROB II strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the second predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	0	86	545
2	130	754	59
1	547	13	0

A0(%): 86.50

TS1 : .79

TS2 : .72

TS3 : .79

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	17	35	42
2	257	361	228
1	68	37	22

A0(%): 44.14

TS1 : .17

TS2 : .39

TS3 : .12

Figure 12c. Cloud amount skill diagram and contingency table results for area 4, TAU-00 (PR+BMD model) for the natural regression strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the third predictor.

SIGNIFICANCE TEST

Null Hypothesis Contingency

Table (Chance)

FORECAST \ OBSERVED	1	2	3
3	44	77	79
2	44	76	79
1	43	76	79

AO(%) 33.17

TS1 .15

TS2 : .22

TS3 .22

Area 2 TAU00

95% CONFIDENCE INTERVAL

AO(%): 29.52 - 37.08

TS1 : .12 - .18

TS2 : .18 - .25

TS3 .19 - .26

Figure 13. Confidence intervals for significance with respect to chance--area 2, TAU-00, cloud amount

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	38	128	298
2	174	331	189
1	16	16	4

AO(%): 54.02

TS1 : .06

TS2 : .39

TS3 : .45

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	29	71	133
2	89	149	100
1	13	9	4

AO(%) 49.41

TS1 : .09

TS2 : .36

TS3 : .39

Figure 14. Contingency table results for the area 2, TAU-00, single-stage regression, EVAR model for cloud amount

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	35	121	284
2	179	339	204
1	14	15	4

AO(%): 53.76

TS1 : .05

TS2 : .41

TS3 : .42

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	28	64	125
2	91	160	108
1	12	5	4

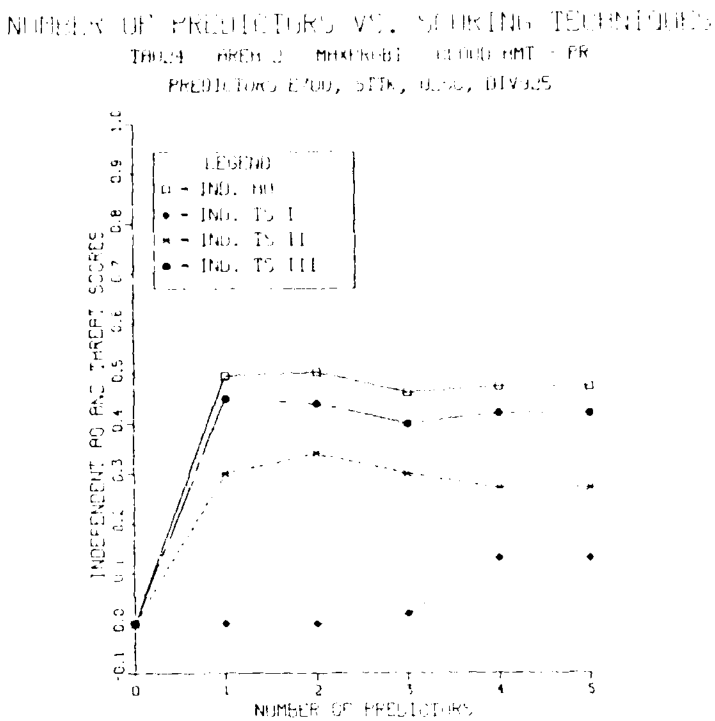
AO(%): 49.02

TS1 .08

TS2 : .38

TS3 : .37

Figure 15. Contingency table results for the area 2, TAU-00, single-stage regression, QUAD model for cloud amount



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	55	143	369
2	60	208	43
1	66	22	16

AO(%)	65.48
TS1	.30
TS2	.44
TS3	.59

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	49	75	145
2	60	102	60
1	0	0	0

AO(%)	50.31
TS1	.0
TS2	.34
TS3	.44

Figure 25a. Cloud amount skill diagram and contingency table results for area 2, TAU-24 (PR model) for the MAXPROB I strategies. The contingency table corresponds to those scores associated with the maximum independent AO achieved at the second predictor.

BASELINE INTERVAL

Baseline model : Area 2 TAU24 BMD-EVAR

FORECAST			
3	30	89	127
2	82	120	68
1	1	8	3
OBSERVED	1	2	3

AO(%)	46.97
TS1	.01
TS2	33
TS3	.40

BASELINE CONFIDENCE INTERVAL

AO(%)	42.97 - 50.97
TS1	.00 - .02
TS2	.32 - .39
TS3	.36 - .44

Figure 24. Confidence intervals for significance with respect to baseline--area 2, TAU-24, cloud amount

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	57	171	269
2	66	113	90
1	82	132	76

AO(%): 43.94

TS1 : .20

TS2 : .20

TS3 : .41

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	30	89	127
2	46	58	37
1	37	70	34

AO(%): 42.05

TS1 : .17

TS2 : .19

TS3 : .40

Figure 23. Contingency table results for the area 2, TAU-24, single-stage regression, MLDC model for cloud amount

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	34	126	226
2	167	278	200
1	4	12	9

AO(%): 48.11

TS1 : .02

TS2 : .36

TS3 : .38

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	21	68	104
2	91	141	91
1	1	8	3

AO(%): 46.59

TS1 : .01

TS2 : .35

TS3 : .36

Figure 22. Contingency table results for the area 2, TAU-24, single-stage regression, QUAD model for cloud amount

DEPENDENT DATA

FORECAST			
3	57	171	269
2	144	233	157
1	4	12	9
OBSERVED	1	2	3

AO(%): 47.92

TS1 : .02

TS2 : .32

TS3 : .41

INDEPENDENT DATA

FORECAST			
3	30	89	127
2	82	120	68
1	1	8	3
OBSERVED	1	2	3

AO(%): 46.97

TS1 : .01

TS2 : .33

TS3 : .40

Figure 21. Contingency table results for the area 2, TAU-24, single-stage regression, EVAR model for cloud amount

SIGNIFICANCE TEST

Null Hypothesis Contingency

Table (Chance)

FORECAST			
3	36	70	77
2	35	70	76
1	35	70	76
	1	2	3
	OBSERVED		

A0(%): 33.39

TS1 .14

TS2 .22

TS3 .23

Area 2 TAU24

95% CONFIDENCE INTERVAL

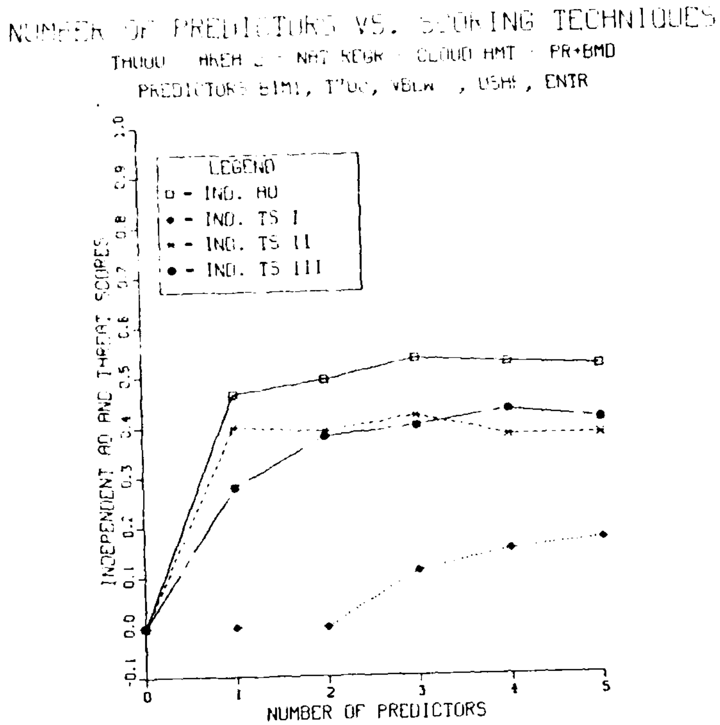
A0(%): 29.43 - 37.35

TS1 : .11 - .17

TS2 : .18 - .25

TS3 : .20 - .27

Figure 20. Confidence intervals for significance with respect to chance--area 2, TAU-24, cloud amount



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	1	43	462
2	46	429	29
1	181	3	0

AO(%): 89.78

TS1 : .78

TS2 : .78

TS3 : .86

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	21	43	120
2	94	183	109
1	16	3	8

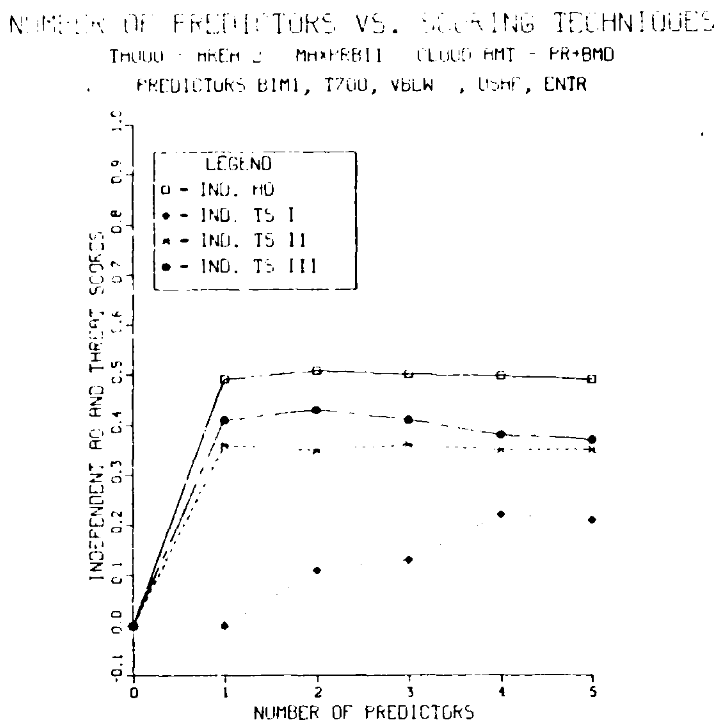
AO(%): 53.43

TS1 : .11

TS2 : .42

TS3 : .40

Figure 19c. Cloud amount skill diagram and contingency table results for area 2, TAU-00 (PR+BMD model) for the natural regression strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the third predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	3	13	436
2	7	434	40
1	218	28	15

AO(%): 91.12

TS1 : .80

TS2 : .83

TS3 : .86

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	33	54	134
2	74	141	86
1	24	34	17

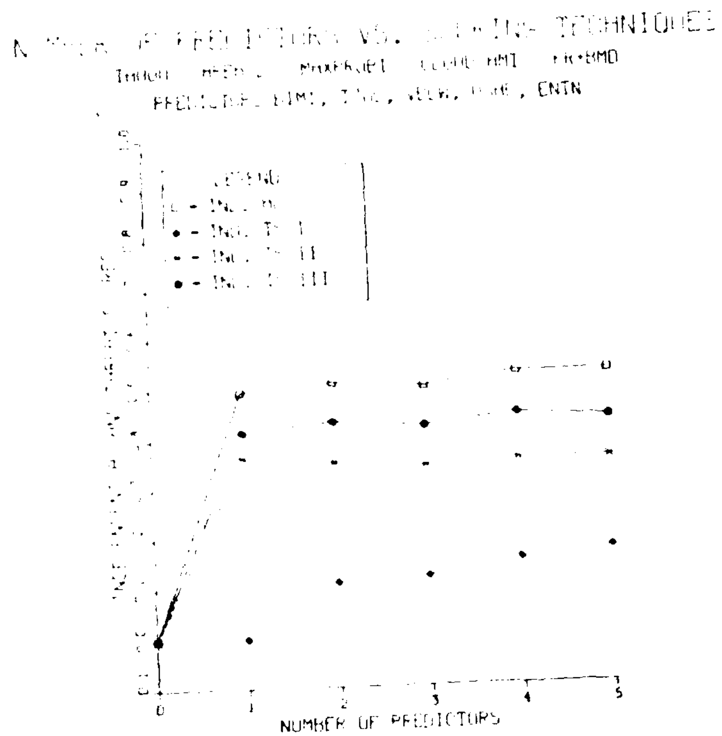
AO(%): 50.08

TS1 : .13

TS2 : .36

TS3 : .41

Figure 19b. Cloud amount skill diagram and contingency table results for area 2, TAU-00 (PR+BMD model) for the MAXPROB II strategies. The contingency table corresponds to those scores associated with the maximum independent AO achieved at the third predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	18	46	481
2	29	426	10
1	181	3	0

A0(%): 91.12

TS1 : .78

TS2 : .83

TS3 : .87

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	39	74	150
2	62	126	68
1	30	29	19

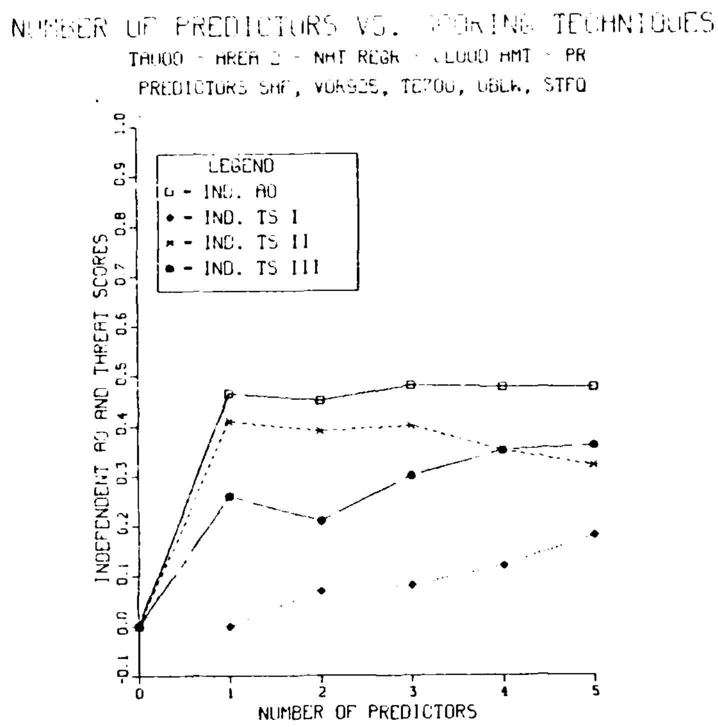
A0(%): 51.26

TS1 : .17

TS2 : .35

TS3 : .43

Figure 19a. Cloud amount skill diagram and contingency table results for area 2, TAU-00 (PR+BMD model) for the MAXPROB I strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the fourth predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	1	51	435
2	79	412	55
1	148	12	1

AO(%): 83.33

TS1 : .61

TS2 : .68

TS3 : .80

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	18	31	87
2	102	189	145
1	11	9	5

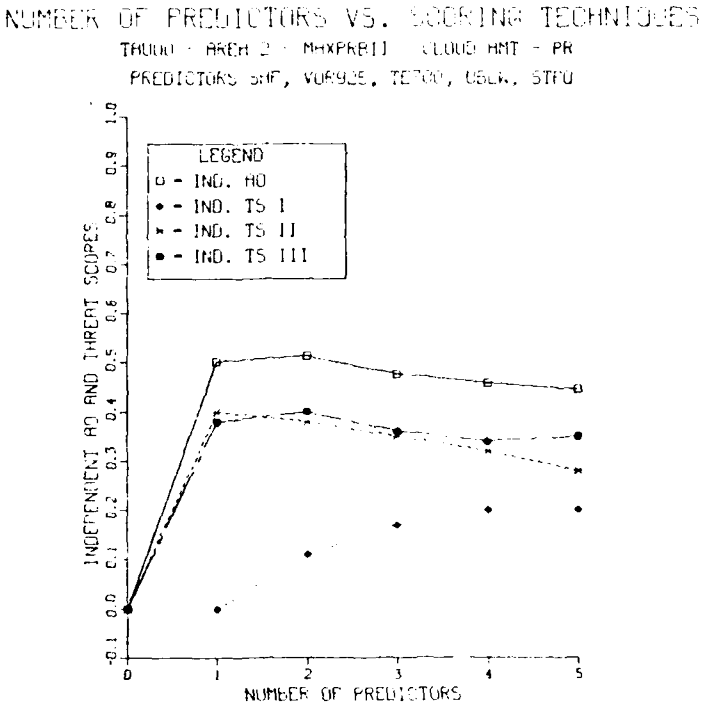
AO(%): 48.07

TS1 .08

TS2 : .40

TS3 : .30

Figure 18c. Cloud amount skill diagram and contingency table results for area 2, TAU-00 (PR model) for the natural regression strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the third predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	7	17	408
2	14	405	61
1	207	53	22

AO(%): 85.43

TS1 : .68

TS2 : .74

TS3 : .79

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	42	65	138
2	73	152	93
1	16	12	6

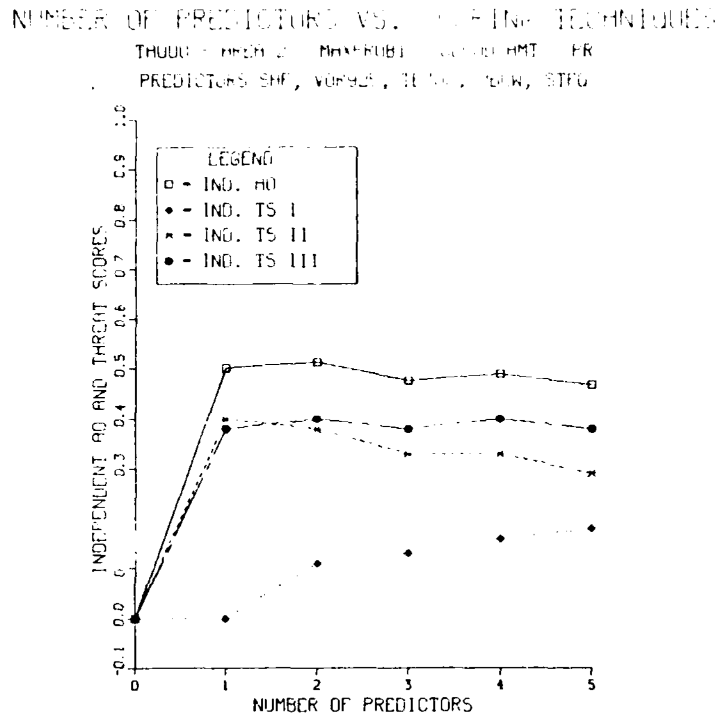
AO(%): 51.26

TS1 .11

TS2 .38

TS3 .40

Figure 18b. Cloud amount skill diagram and contingency table results for area 2, TAU-00 (PR model) for the MAXPROB II strategies. The contingency table corresponds to those scores associated with the maximum independent AO achieved at the second predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	26	64	468
2	47	397	19
1	155	14	4

A0(%): 85.43

TS1 : .63

TS2 : .73

TS3 : .81

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	42	65	138
2	73	152	93
1	16	12	6

A0(%): 51.26

TS1 : .11

TS2 : .38

TS3 : .40

Figure 18a. Cloud amount skill diagram and contingency table results for area 2, TAU-00 (PR model) for the MAXPROB I strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the second predictor.

BASELINE INTERVAL

Baseline model : Area TAU00 BMD-EVAR

FORECAST			
3	29	71	133
2	89	149	100
1	13	9	4
	1	2	3
	OBSERVED		

A0(%)	49.41
TS1	.09
TS2	.36
TS3	.39

BASELINE CONFIDENCE INTERVAL

A0(%)	45.40 - 53.42
TS1	.07 - .11
TS2	.32 - .39
TS3	.36 - .44

Figure 17. Confidence intervals for significance with respect to baseline--area 2, TAU-00, cloud amount

DEPENDENT DATA

FORECAST	Observed 1	Observed 2	Observed 3
3	38	128	298
2	69	199	114
1	121	148	79

OBSERVED

AO(%): 51.76

TS1 : .27

TS2 : .30

TS3 : .45

INDEPENDENT DATA

FORECAST	Observed 1	Observed 2	Observed 3
3	29	71	133
2	37	85	60
1	65	73	44

OBSERVED

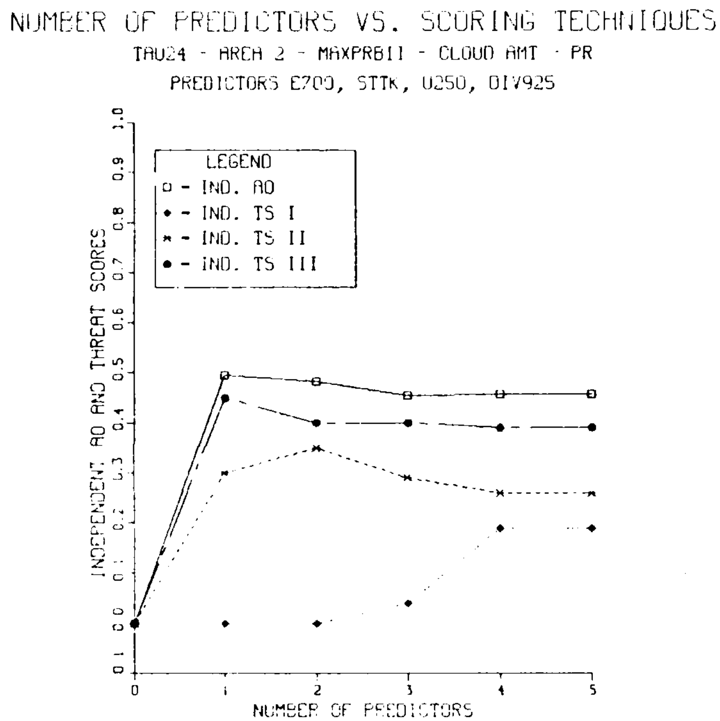
AO(%): 47.40

TS1 : .26

TS2 : .26

TS3 : .39

Figure 16. Contingency table results for the area 2, TAU-00, single-stage regression, MLDC model for cloud amount



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	30	81	300
2	46	238	88
1	105	54	40

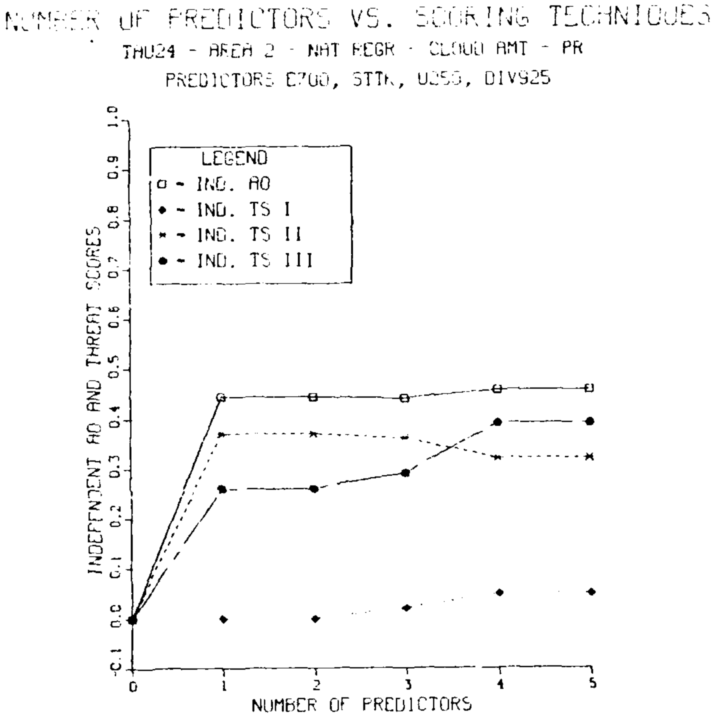
AO(%): 65.48
TS1 : .38
TS2 : .47
TS3 : .56

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	61	97	163
2	48	80	42
1	0	0	0

AO(%): 49.49
TS1 : .0
TS2 : .30
TS3 : .45

Figure 25b. Cloud amount skill diagram and contingency table results for area 2, TAU-24 (PR model) for the MAXPROB II strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the first predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	5	91	277
2	137	276	149
1	39	6	2

AO(%): 60.29

TS1 : .21

TS2 : .42

TS3 : .53

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	23	56	110
2	80	108	85
1	6	13	10

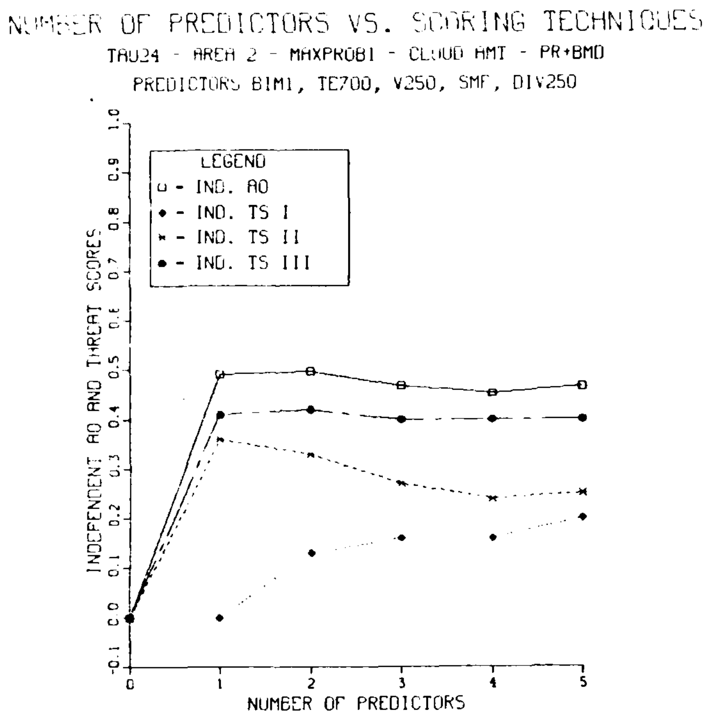
AO(%): 45.62

TS1 : .05

TS2 : .32

TS3 : .39

Figure 25c. Cloud amount skill diagram and contingency table results for area 2, TAU-24 (PR model) for the natural regression strategies. The contingency table corresponds to those scores associated with the maximum independent AO achieved at the fourth predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	16	39	422
2	16	331	6
1	149	3	0

AO(%): 91.85

TS1 : .81

TS2 : .84

TS3 : .87

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	29	66	125
2	65	104	77
1	15	7	3

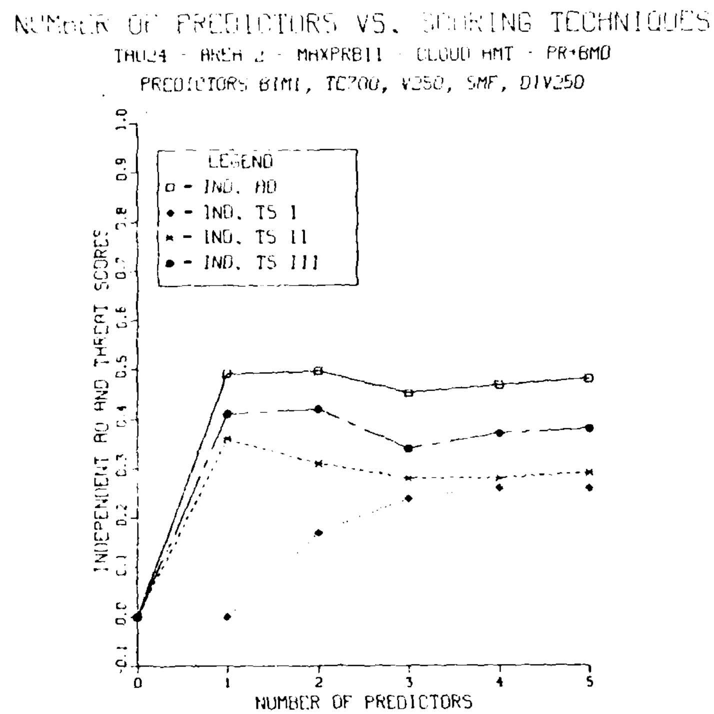
AO(%): 49.69

TS1 : .13

TS2 : .33

TS3 : .42

Figure 26a. Cloud amount skill diagram and contingency table results for area 2, TAU-24 (PR+BMD model) for the MAXPROB I strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the second predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	2	1	376
2	1	348	38
1	178	24	14

AO(%): **91.85**

TS1 : **.81**

TS2 : **.84**

TS3 : **.87**

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	29	66	125
2	57	96	72
1	23	15	8

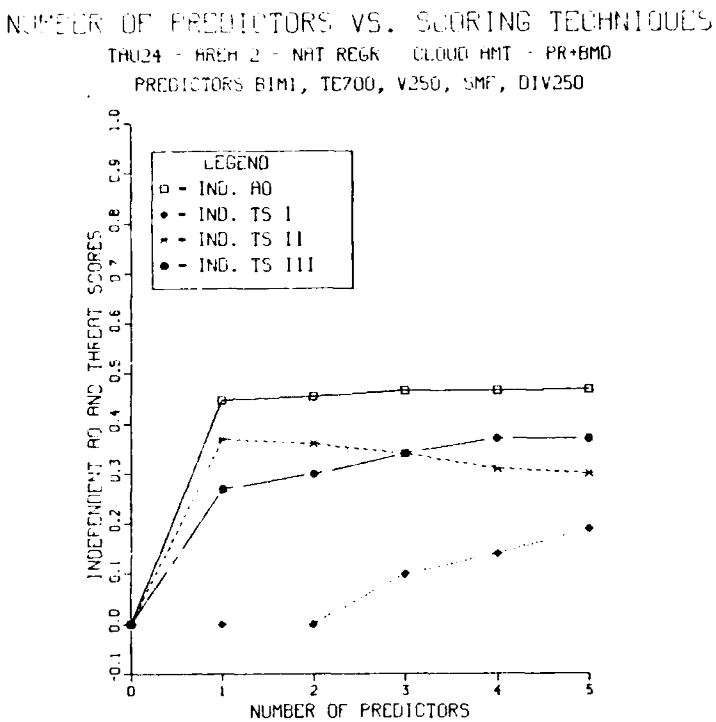
AO(%): **49.69**

TS1 : **.17**

TS2 : **.31**

TS3 : **.42**

Figure 26b. Cloud amount skill diagram and contingency table results for area 2, TAU-24 (PR+BMD model) for the MAXPROB II strategies. The contingency table corresponds to those scores associated with the maximum independent AO achieved at the second predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	1	33	406
2	31	337	22
1	149	3	0

AO(%): 90.84

TS1 : .81

TS2 : .79

TS3 : .88

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	23	61	108
2	67	102	88
1	19	14	9

AO(%). 46.64

TS1 : .14

TS2 : .31

TS3 : .37

Figure 26c. Cloud amount skill diagram and contingency table results for area 2, TAU-24 (PR+BMD model) for the natural regression strategies. The contingency table corresponds to those scores associated with the maximum independent AO achieved at the fifth predictor.

SIGNIFICANCE TEST

Null Hypothesis Contingency

Table (Chance)

FORECAST \ OBSERVED	1	2	3
3	40	69	80
2	40	69	79
1	40	68	79

AO(%): 33.51

TS1 .15

TS2 : .21

TS3 : .23

Area 2 TAU 48

95% CONFIDENCE INTERVAL

AO(%): 29.30 - 37.11

TS1 .12 - .18

TS2 .18 - .20

TS3 .20 - .26

Figure 27. Confidence intervals for significance with respect to chance--area 2, TAU-48, cloud amount

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	59	171	272
2	159	284	176
1	16	21	17

AO(%): 48.68

TS1 : .06

TS2 : .35

TS3 : .39

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	41	93	134
2	75	126	91
1	6	12	9

AO(%): 45.32

TS1 : .04

TS2 : .32

TS3 : .36

Figure 28. Contingency table results for the area 2, TAU-48, single-stage regression, EVAR model for cloud amount

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	59	171	272
2	113	206	136
1	62	99	57

A0(%): 45.96

TS1 : .16

TS2 : .28

TS3 : .39

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	41	93	134
2	51	88	69
1	30	50	31

A0(%): 42.93

TS1 : .15

TS2 : .25

TS3 : .36

Figure 29. Contingency table results for the area 2, TAU-48, single-stage regression, MLDC model for cloud amount

BASELINE INTERVAL

Baseline model : Area 2 TAU48 BMD-EVAR

FORECAST \ OBSERVED	1	2	3
3	41	93	134
2	75	126	91
1	6	12	9

A0(%) 45.32

TS1 .04

TS2 32

TS3 .36

BASELINE CONFIDENCE INTERVAL

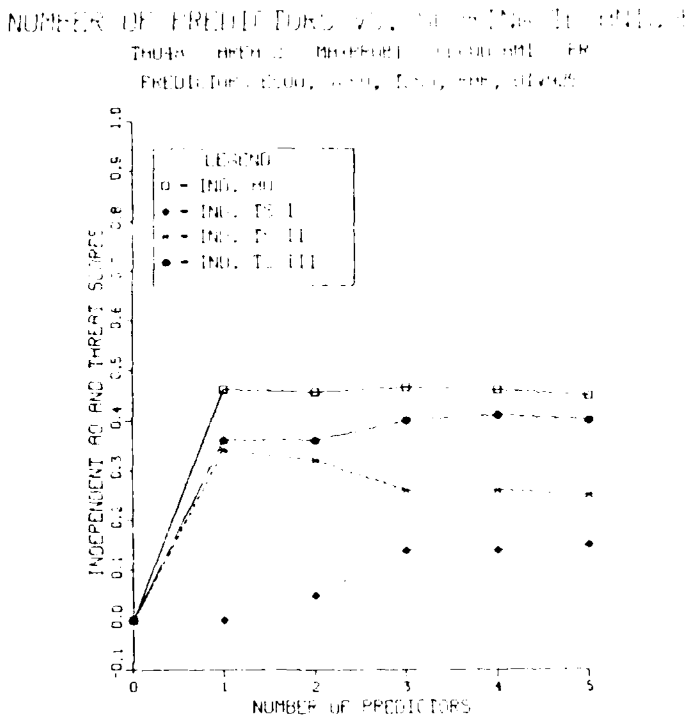
A0(%) **41.31 - 49.22**

TS1 : **.02 - .06**

TS2 : **.28 - .36**

TS3 **.32 - .40**

Figure 30. Confidence intervals for significance with respect to baseline--area 2, TAU-48, cloud amount



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	14	39	465
2	15	411	9
1	179	0	2

A0(%): 93.03

TS1 : .85

TS2 : .87

TS3 : .88

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	50	98	157
2	48	85	75
1	21	23	9

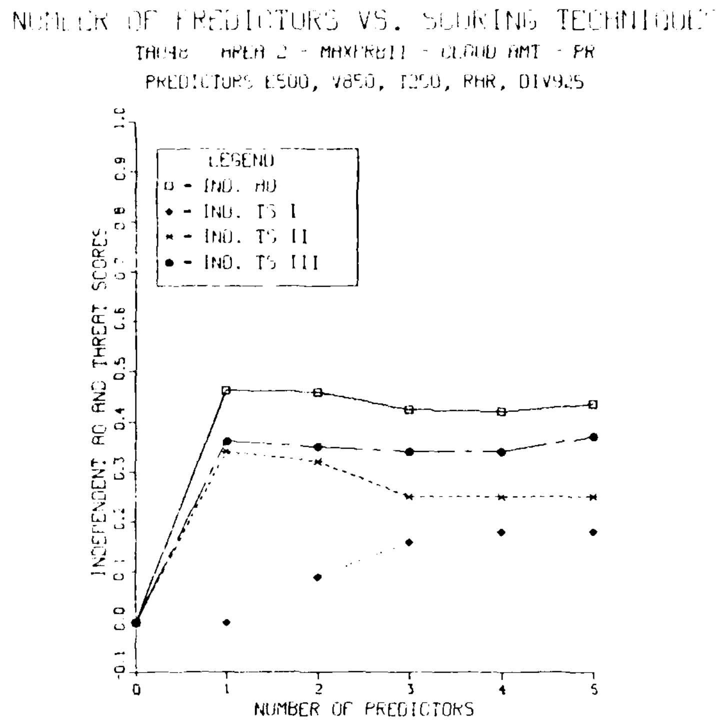
A0(%): 46.47

TS1 : .14

TS2 : .26

TS3 : .40

Figure 31a. Cloud amount skill diagram and contingency table results for area 2, TAU-48 (PR model) for the MAXPROB I strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the third predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	1	7	424
2	5	429	37
1	202	14	15

A0(%): 93.03

TS1 : .85

TS2 .87

TS3 : .88

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	48	73	128
2	58	118	103
1	13	15	10

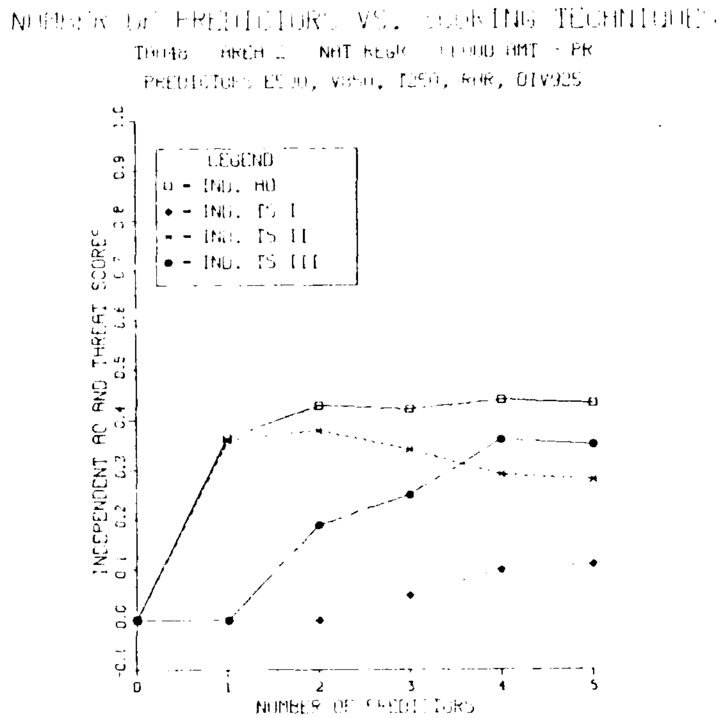
A0(%): 45.76

TS1 .09

TS2 : .32

TS3 : .35

Figure 31b. Cloud amount skill diagram and contingency table results for area 2, TAU-48 (PR model) for the MAXPROB II strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the second predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	0	35	450
2	33	415	26
1	175	0	0

AO(%) 91.71

TS1 : .84

TS2 : .82

TS3 : .88

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	39	64	123
2	62	108	99
1	18	34	19

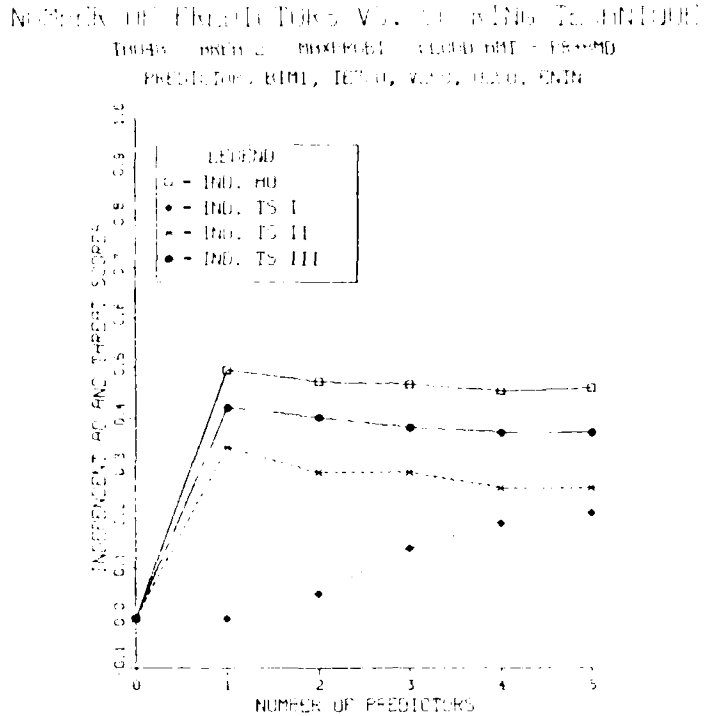
AO(%): 43.99

TS1 : .10

TS2 : .29

TS3 : .36

Figure 31c. Cloud amount skill diagram and contingency table results for area 2, TAU-48 (PR model) for the natural regression strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the fourth predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	11	37	467
2	18	409	7
1	179	4	2

A0(%): 93.03

TS1: .84

TS2: .86

TS3: .89

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	43	81	155
2	76	125	86
1	0	0	0

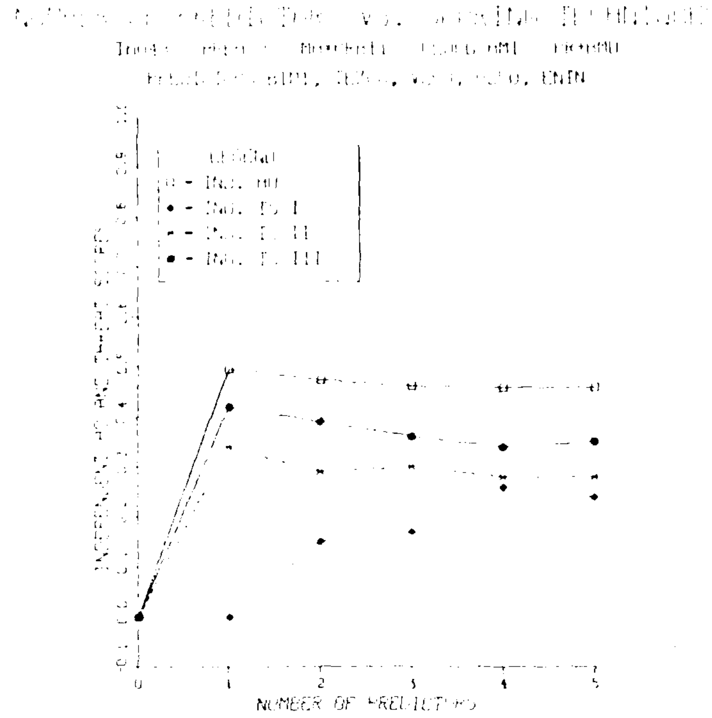
A0(%): 49.47

TS1: .0

TS2: .34

TS3: .42

Figure 32a. Cloud amount skill diagram and contingency table results for area 2, TAU-48 (PR+BMD model) for the MAXPROB I strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the first predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	2	7	428
2	7	428	37
1	199	15	11

AO(%) 93.03

TS1 .85

TS2 .87

TS3 .88

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	43	81	155
2	76	125	86
1	0	0	0

AO(%) 49.47

TS1 .0

TS2 .34

TS3 .42

Figure 32b. Cloud amount skill diagram and contingency table results for area 2, TAU-48 (PR+BMD model) for the MAXPROB II strategies. The contingency table corresponds to those scores associated with the maximum independent AO achieved at the first predictor.

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	5	40	242
2	173	408	108
1	137	71	10

AO(%): 65.91

TS1 : .35

TS2 : .51

TS3 : .60

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	4	20	137
2	73	204	59
1	62	36	2

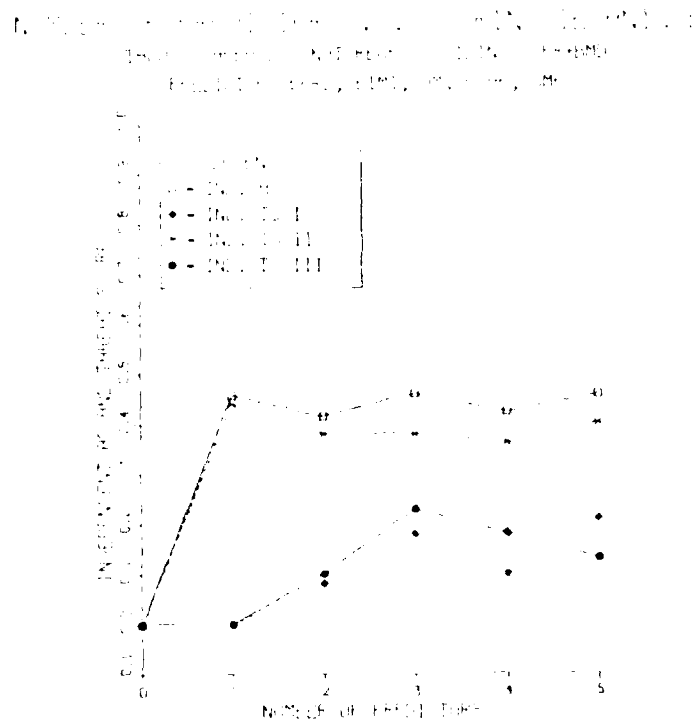
AO(%): 67.50

TS1 : .35

TS2 : .52

TS3 : .62

Figure 42. Contingency table results for the area 2, TAU-00, single-stage regression, using cloud amount as a predictor EVAR model for ceiling



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	3	38	86
2	232	456	266
1	80	25	8

A0(%) 52.09

TS1 .23

TS2 .45

TS3 .21

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	18	66	64
2	92	182	128
1	29	12	6

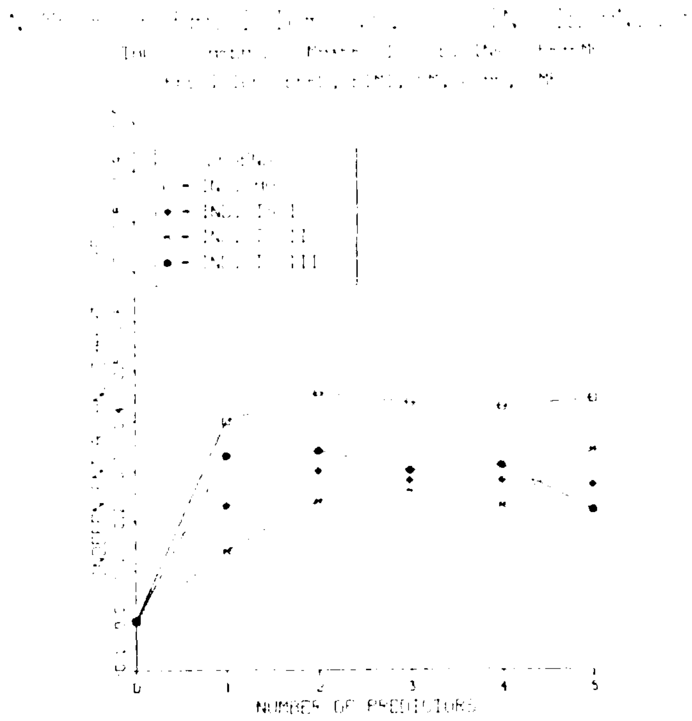
A0(%) 46.06

TS1 .18

TS2 .38

TS3 .23

Figure 41c. Ceiling skill diagram and contingency table results for area 2, TAU-00 (PR+BMD model) for the natural regression strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the third predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	39	58	139
2	95	369	162
1	181	92	59

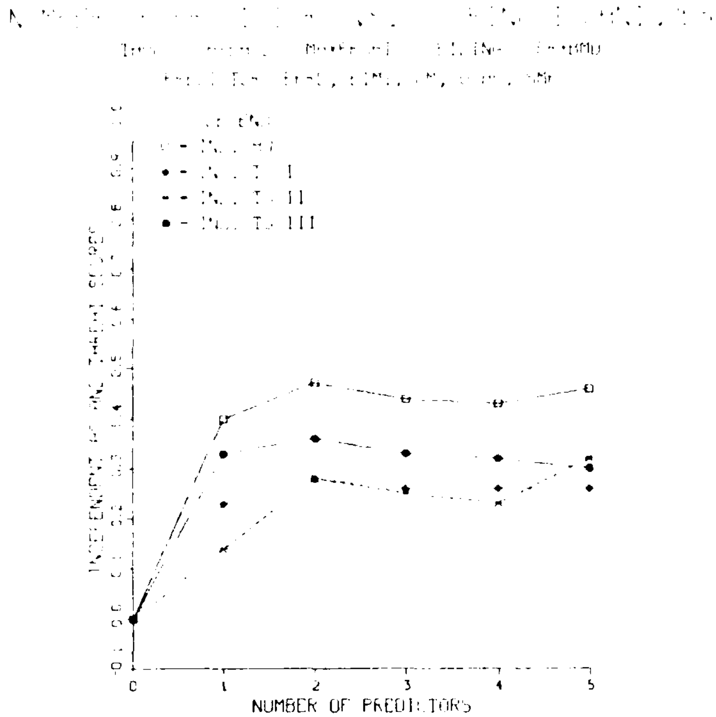
A0(%)	57.71
TS1	.39
TS2	.48
TS3	.30

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	38	121	122
2	33	83	48
1	68	56	28

A0(%)	45.73
TS1	.30
TS2	.24
TS3	.34

Figure 41b. Ceiling skill diagram and contingency table results for area 2, TAU-00 (PR+BMD model) for the MAXPROB II strategies. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the second predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	75	114	208
2	107	348	124
1	133	57	28

AO(%) 57.71

TS1 .33

TS2 .46

TS3 .38

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	43	129	132
2	45	99	52
1	51	32	14

AO(%) 47.24

TS1 .28

TS2 .28

TS3 .36

Figure 41a. Ceiling skill diagram and contingency table results for area 2, TAU-00 (PR+BMD model) for the MAXPROB I strategies. The contingency table corresponds to those scores associated with the maximum independent AO achieved at the second predictor.

BASELINE INTERVAL

Baseline model : Area 2 TAU00 BMD-MLDC

FORECAST \ OBSERVED	1	2	3
3	6	28	34
2	89	194	145
1	44	38	19

A0(%)	45.56
TS1	.22
TS2	.39
TS3	.15

BASELINE CONFIDENCE INTERVAL

A0(%)	41.56 - 49.56
TS1	.19 - .25
TS2	.35 - .42
TS3	.12 - .18

Figure 40. Confidence intervals for significance with respect to baseline--area 2, TAU-00, ceiling

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	12	60	57
2	197	398	269
1	106	61	34

A0(%): 46.98

TS1 : .26

TS2 : .40

TS3 : .13

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	6	28	34
2	89	194	145
1	44	38	19

A0(%): 45.56

TS1 .22

TS2 .39

TS3 .15

Figure 39. Contingency table results for the area 2, TAU-00, single-stage regression, MLDC model for ceiling

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	0	0	0
2	209	458	326
1	106	61	34

AO(%): 47.24

TS1 : .26

TS2 : .43

TS3 : .0

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	0	0	0
2	95	222	179
1	44	38	19

AO(%): 44.56

TS1 : .22

TS2 : .42

TS3 : .0

Figure 38. Contingency table results for the area 2, TAU-00, single-stage regression, EVAR model for ceiling

SIGNIFICANCE TEST

Null Hypothesis Contingency

Table (Chance)

FORECAST		OBSERVED 1	OBSERVED 2	OBSERVED 3
	3	47	87	66
	2	46	87	66
	1	46	86	66

A0(%)	33.33
TS1	.16
TS2	.23
TS3	.20

Area 2 TAU00

95% CONFIDENCE INTERVAL

A0(%)	29.23 - 36.77
TS1	.13 - .19
TS2	.20 - .27
TS3	.17 - .23

Figure 37. Confidence intervals for significance with respect to chance--area 2, TAU-00, ceiling

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	46	107	264
2	176	365	223
1	6	3	4

AO(%): 53.18

TS1 : .03

TS2 : .42

TS3 : .41

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	26	56	124
2	103	173	110
1	2	0	3

AO(%): 50.08

TS1 : .01

TS2 : .39

TS3 : .39

Figure 36b. Contingency table results for the area 2, TAU-00, two-stage regression, predictors chosen by combination of clustering and separability techniques for cloud amount

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	48	130	290
2	169	342	194
1	11	3	7

AO(%): 53.85

TS1 : .05

TS2 : .41

TS3 : .43

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	34	67	133
2	90	162	100
1	7	0	4

AO(%): 50.59

TS1 : .05

TS2 : .39

TS3 : .39

Figure 36a. Contingency table results for the area 2, TAU-00, two-stage regression, EVAR, predictors chosen by combination of clustering and separability techniques for cloud amount

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	48	154	270
2	180	321	221
1	0	0	0

A0(%): 49.50

TS1 : .0

TS2 : .37

TS3 : .39

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	39	75	109
2	92	154	128
1	0	0	0

A0(%): 44.05

TS1 : .0

TS2 : .34

TS3 : .31

Figure 35. Contingency table results for the area 2, TAU-00, single-stage regression, predictors chosen by combination of clustering and separability techniques for cloud amount

DEPENDENT DATA

FORECAST	Observed 1	Observed 2	Observed 3
3	44	125	290
2	184	350	201
1	0	0	0

OBSERVED

AO(%): 53.60

TS1 : .0

TS2 : .41

TS3 : .44

INDEPENDENT DATA

FORECAST	Observed 1	Observed 2	Observed 3
3	23	68	131
2	108	161	106
1	0	0	0

OBSERVED

AO(%): 48.91

TS1 : .0

TS2 : .36

TS3 : .40

Figure 34. Contingency table results for the area 2, TAU-00, single-stage regression, predictors chosen by highest measures of separability for category II versus III for cloud amount

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	58	155	287
2	170	320	204
1	0	0	0

AO(%): 50.84

TS1 : .0

TS2 : .38

TS3 : .41

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	31	76	118
2	100	153	119
1	0	0	0

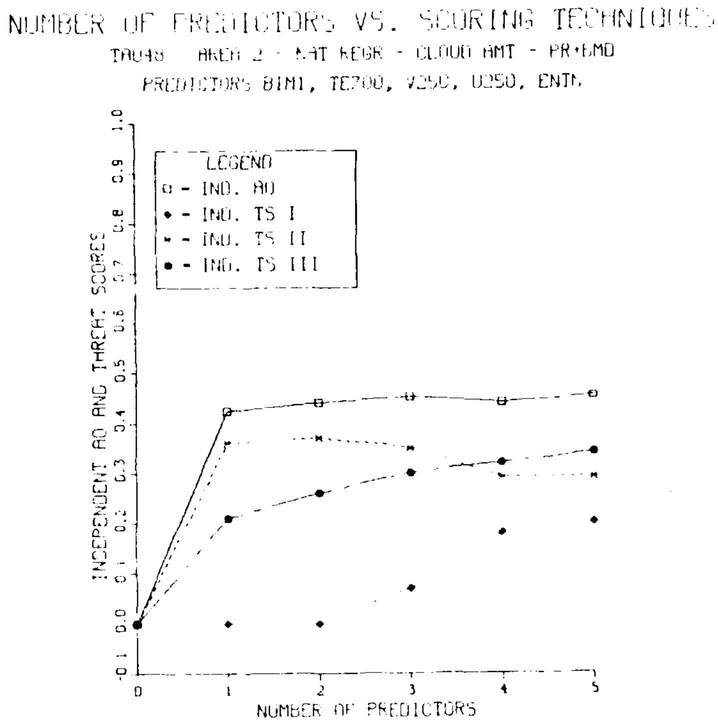
AO(%) 45.39

TS1 : .0

TS2 : .34

TS3 : .34

Figure 33. Contingency table results for the area 2, TAU-00, single-stage regression, predictors chosen by highest measures of separability for category I versus II for cloud amount



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	0	36	452
2	34	412	24
1	174	2	0

AO(%): 91.53

TS1 : .83

TS2 : .81

TS3 : .88

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	32	78	118
2	55	106	105
1	32	22	18

AO(%): 45.23

TS1 : .20

TS2 : .29

TS3 : .34

Figure 32c. Cloud amount skill diagram and contingency table results for area 2, TAU-48 (PR+BMD model) for the natural regression strategies. The contingency table corresponds to those scores associated with the maximum independent AO achieved at the fifth predictor.

DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	5	39	240
2	182	414	110
1	128	66	10

AO(%): 65.49

TS1 : .33

TS2 : .51

TS3 : .59

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	4	19	137
2	76	208	59
1	59	33	2

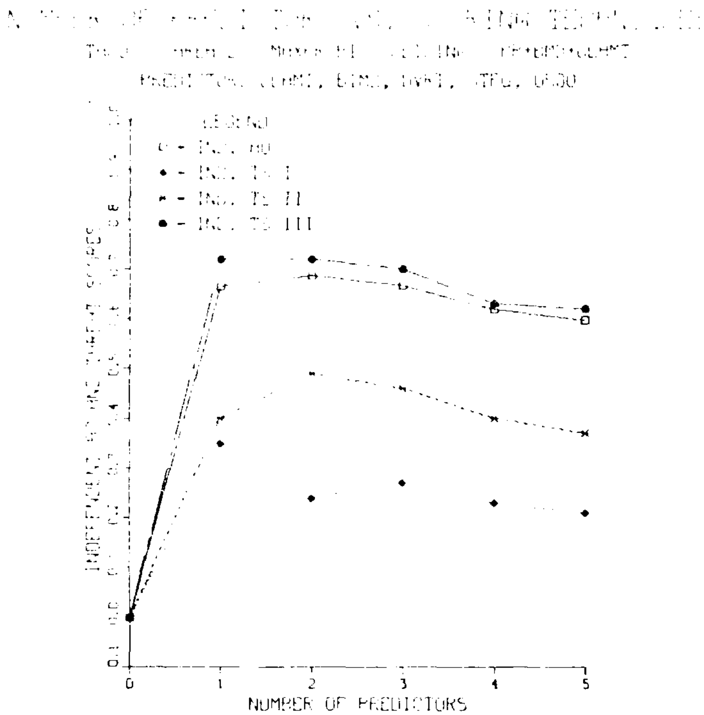
AO(%): 67.67

TS1 : .34

TS2 : .53

TS3 : .62

Figure 43. Contingency table results for the area 2, TAU-00, single-stage regression, using cloud amount as a predictor QUAD model for ceiling



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	29	48	843
2	89	722	6
1	378	14	5

A0(%) 91.05

TS1 : .73

TS2 : .82

TS3 : .91

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	17	52	364
2	146	299	65
1	69	45	10

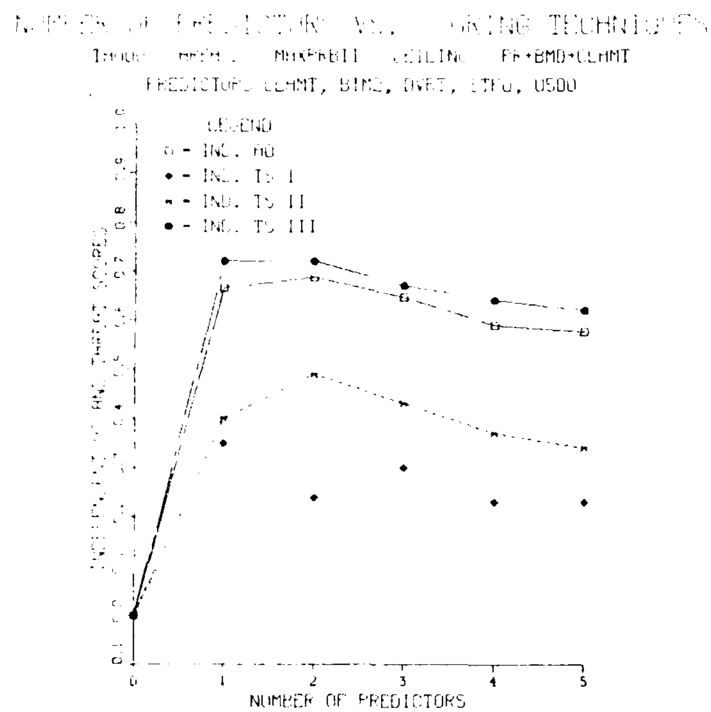
A0(%) 68.60

TS1 : .24

TS2 : .49

TS3 : .72

Figure 44a. Ceiling skill diagram and contingency table results for area 2, TAU-00 (PR+BMD model) for the MAXPROB I strategies, including cloud amount as a predictor. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the second predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	1	5	779
2	13	682	41
1	482	97	34

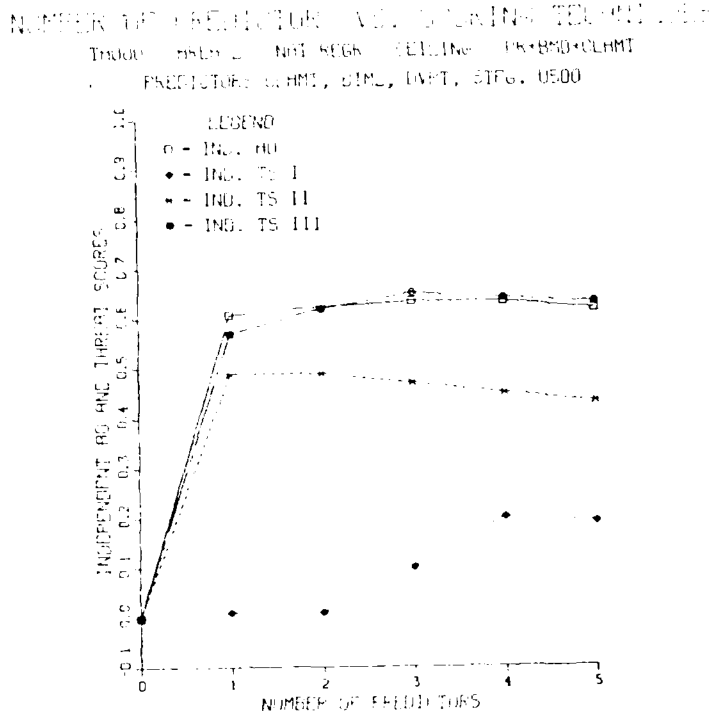
AO(%): 91.05
TS1 : .77
TS2 : .81
TS3 : .91

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	17	52	364
2	146	299	65
1	69	45	10

AO(%): 68.60
TS1 : .24
TS2 : .49
TS3 : .72

Figure 44b. Ceiling skill diagram and contingency table results for area 2, TAU-00 (PR+BMD model) for the MAXPROB II strategies, including cloud amount as a predictor. The contingency table corresponds to those scores associated with the maximum independent AO achieved at the second predictor.



DEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	0	40	813
2	130	733	41
1	366	11	0

A0(%) 89.60

TS1 : .72

TS2 : .77

TS3 : .91

INDEPENDENT DATA

FORECAST \ OBSERVED	1	2	3
3	21	42	321
2	149	293	105
1	62	61	13

A0(%): 63.36

TS1 : .20

TS2 : .45

TS3 : .64

Figure 44c. Ceiling skill diagram and contingency table results for area 2, TAU-00 (PR+BMD model) for the natural regression strategies, including cloud amount as a predictor. The contingency table corresponds to those scores associated with the maximum independent A0 achieved at the fourth predictor.

LIST OF REFERENCES

Aldinger, W.T., 1979: Experiments on Estimating Open Ocean Visibilities Using Model Output Statistics. M.S. Thesis (R.J. Renard, advisor), Dept. of Meteorology, Naval Postgraduate School, Monterey, CA, 81 pp.

Best, D.L., and Pryor, S.P., 1983: Air Weather Service Model Output Statistics System Project Report. AFGWC/PR-83/001, United States Air Force Air Weather Service (MAC), Air Force Global Weather Central, Offutt AFB, NE, 95 pp.

Cooley, Duane S., 1978: The Use of Model Output Statistics for Predicting Ceiling, Visibility and Cloud Amount, NWS Technical Procedures Bulletin No. 234, NOAA, U.S. Department of Commerce, p. 22.

Department of the Navy, 1979: CV NATOPS MANUAL, Office of the Chief of Naval Operations, Washington, D.C., 156 pp.

Diunizio, M., 1984: An Evaluation of Discretized Conditional Probability and Linear Regression Theshold Techniques in Model Output Statistics Forecasting of Visibility over the North Atlantic Ocean. M.S. Thesis (R.J. Renard, advisor), Dept. of Meteorology, Naval Postgraduate School, Monterey, CA, 227 pp.

Glahn, H.R., 1983: MOS Support for Military Locations From the Techniques Development Laboratory. TDL Office Note 83-9, National Weather Service, NOAA, U.S. Department of Commerce, 10 pp.

________, and Lowry, D.A., 1972: The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. J. Appl. Meteor., 11, pp. 1203-1211.

Godfrey, R.S. and P.R. Lowe, 1979: An Application of Model Output Statistics to Forecasting the Occurrence of the Levante Wind. Preprints, Sixth Conference on Probability and Statistics in Atmospheric Sciences, Banff, Alta., Amer. Meteor. Soc., Boston, MA, pp. 83-86.

Hand, D.J., 1981: Discrimination and Classification, John Wiley and Sons, New York, N.H., pp. 131-159.

Karl, M.L., 1984: Experiments in Forecasting Atmospheric Marine Horizontal Visibility using Model OUtput Statistis with Conditional Probabilities of Discretized Parameters. M.S. Thesis (R.J. Renard, advisor), Dept. of Meteorology, Naval Postgraduate School, Monterey, CA, 165 pp.

Klein, W.H., 1981: Design of a MOS System for the Navy. Final Report. NEPRF Contract No. N00228-80-C-LV22, Intercon Weather Consultants, Inc., Camp Springs, MD.

Koziara, M.C., R.J. Renard and W.J. Thompson, 1983: Estimating Marine Fog Probability Using a Model Output Statistics Scheme. Monthly Weather Review, 111, pp. 2333-2340.

Lewit, H.L., 1980: SOCAL MOS Ceiling/Visibility Forecast Algorithms. Task 3 Final Report: Forecast Algorithm Development, and Task IV, Section I: Test of Forecast Algorithms Employing Dependent Data. NEPRD Contract No. N00228-78-C-3289, Ocean Data Systems, Inc., Monterey, CA.

Lowe, P., 1984a: The Use of Decision Theory for Determining Thresholds for Categorical Forecasts, unpublished manuscript, Naval Environmental Prediction Research Facility, Monterey, CA.

________, 1984b: The Use of Multi-Variate Statistics for Defining Homogeneous Atmospheric REgions Over the North Atlantic Ocean, unpublished manuscript, Naval Environmental Prediction Research Facility, Monterey, CA.

________, 1984c: The Use of Adjusted Scores and Significance Tests in the Verification of Categorical Forecasts, unpublished manuscript, Naval Environmental Prediction Research Facility, Monterey, CA.

Maglaras, George, 1984: Alaskan Ceiling Height, Visibility, and Obstructions to Vision Guidance, NWS Technical Procedures Bulletin, No. 334, NOAA, U.S. Department of Commerce, p. 7.

Miller, I., and Freund, J.E., 1977: Probability and Statistics for Engineers, 2nd. Ed., Prentice-Hall, Inc., Englewood Cliffs, NJ, 528 pp.

Naval Environmental Prediction Research Facility, Monterey, CA, 1982: MOS Forecasts for U.S. Navy and Marine Corps CONUS Locations--User's Manual, NAVENPREDRSHFAC Document No. 7W0513-Um-07, Commander, Naval Oceanography Command, Bay St. Louis, MS, 13 pp.

Nelson, T.S., 1972: Numerical-Statistical Prediction of Visibility at Sea. M.S. Thesis (R.J. Renard, advisor), Dept. of Meteorology, Naval Postgraduate School, Monterey, CA, 33 pp.

Preisendorfer, R.W., 1983a: Proposed Studies of Some Basic Marine Atmospheric Visibility Prediction Schemes Using Model Output Statistics, unpublished manuscript, Department of Meteorology, Naval Postgraduate School, Monterey, CA, 28 pp.

________, 1983b: Maximum-Probability and Natural-Regression Prediction Strategies, unpublished manuscript, Department of Meteorology, Naval Postgraduate School, Monterey, CA, 10 pp.

________, 1983c: Tests for Functional Dependence of Predictors, unpublished manuscript, Department of Meteorology, Naval Postgraduate School, Monterey, CA, 5 pp.

________, 1984: Update of MAXPROB MOS Prediction Method, unpublished manuscript, Department of Meteorology, Naval Postgraduate School, Monterey, CA, 3 pp.

Renard, R.J. and W.T. Thompson, 1984: Estimating Visibility Over the North Pacific Ocean Using Model Output Statistics. National Weather Digest, Vol. 9, No. 2, pp. 18-25.

Schramm, W.G., 1966: Analysis and Prediction of Visibility at Sea. M.S. Thesis (R.J. Renard, advisor), Dept. of Meteorology, Naval Postgraduate School, Monterey, CA, 224 pp.

Selsor, H.D., 1980: Further Experiments Using a Model Output Statistics Method in Estimating Open Ocean Visibility. M.S. Thesis (R.J. Renard, advisor), Dept. of Meteorology, Naval Postgraduate School, Monterey, CA, 121 pp.

University of California, 1983: BMDP Statistical Software, 1983 Edition, Department of Biomathematics, University of California at Los Angeles, University of California Press, 726 pp.

Yavorsky, P.G., 1980: Experiments Concerning Categorical Forecasts of Open-Ocean Visibility Using Model Output Statistics. M.S. Thesis (R.J. Renard, advisor), Dept. of Meteorology, Naval Postgraduate School, Monterey, CA, 87 pp.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, CA 93943	2
3.	Meteorology Reference Center, Code 63 Department of Meteorology Naval Postgraduate School Monterey, CA 93943	1
4.	Professor Robert J. Renard, Code 63Rd Chairman, Department of Meteorology Naval Postgraduate School Monterey, CA 93943	6
5.	Chairman (Code 68Mr) Department of Oceanography Naval Postgraduate School Monterey, CA 93943	1
6.	Dr. Rudolph W. Preisendorfer NOAA/PMEL/R/E/PM Bin 015700 Bldg. 3 7600 Snad Point Way, N.E. Seattle, WA 98115-0070	1
7.	Mr. Paul Lowe Naval Environmental Prediction Research Facility Monterey, CA 93940	1
8.	Dr. Robert Godfrey Naval Environmental Prediction Research Facility Monterey, CA 93940	1
9.	Lt. Mark Diunizio 10855 Charbono Point San Diego, CA 92131	1
10.	Director Naval Oceanography Division Naval Observatory 34th and Massachusetts Avenue NW Washington, DC 20390	1

11. Commander 1
Naval Oceanography Command
NSTL STation
Bay St. Louis, MS 39522

12. Commanding Officer 1
Naval Oceanographic Office
NSTL Station
Bay St. Louis, MS 39522

13. Commanding Officer 1
Fleet Numerical Oceanography Center
Monterey, CA 93940

14. Commanding Officer 1
Naval Ocean Research and Development Activity
NSTL STation
Bay St. Louis, MS 39522

15. Commanding Officer 1
Naval Environmental Prediction Research Facility
Monterey, CA 93940

16. Chairman, Oceanography Department 1
U.S. Naval Academy
Annapolis, MD 21402

17. Chief of Naval Research 1
800 N. Quincy Street
Arlington, VA 22217

18. Office of Naval Research (Code 480) 1
Naval Ocean Research and Development Activity
NSTL Station
Bay St. Louis, MS 39522

19. Commander (Air-370) 1
Naval Air Systems Command
Washington, D.C. 20360

20. Chief, Ocean Services Division 1
National Oceanic and Atmospheric Administration
8060 Thirteenth Street
Silver Spring, MD 20910

21. Dr. Alan Weinstein 1
Leader, Code 422
Ocean Sciences Division
Office of Naval Research
Arlington, VA 22217

22. Lcdr Michael L. Karl 1
USS Peleliu (LHA-5)
FPO San Francisco, CA 96624

23. Lcdr Mike Wooster 3
217 ADA #22
Mountain View, CA 94043

24. Lcdr Kris Elias 1
SMC #1542
Naval Postgraduate School
Monterey, CA 93943

25. Mr. Gil Ross, Met 09 1
Meteorological Office
Bracknell, Berkshire
England

26. Chief, Technical Procedures Branch 1
Meteorological Services Division
National Oceanic and Atmospheric
Administration
National Weather Service
Silver Spring, MD 20910

27. Chief, Technical Services Division 1
United States Air Force
Air Weather Service (MAC)
Air Force Global Weather Central
Offutt AFB, NB 68113

28. Lt Steve Fatjo 1
SMC #2412
Naval Postgraduate School
Monterey, CA 93943

END

FILMED

8-85

DTIC